Heath–Jarrow–Morton models with jumps

by

Mesias Alfeus

Thesis presented in fulfilment of the requirements for the degree of Master of Science in Financial Mathematics in the Faculty of Science at the University of Stellenbosch.

Supervisor: Dr P.W. Ouwehand

March 2015

Stellenbosch University https://scholar.sun.ac.za

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: ....

M. Alfeus

Date: .... November 6, 2015

Copyright © 2015 Stellenbosch University
All rights reserved.

Stellenbosch University https://scholar.sun.ac.za

Abstract

The standard-Heath–Jarrow–Morton (HJM) framework is well-known for its application to pricing and hedging interest rate derivatives. This study implemented the extended HJM framework introduced by Eberlein and Raible (1999), in which a Brownian motion (BM) is replaced by a wide class of processes with jumps. In particular, the HJM driven by the generalised hyperbolic processes was studied. This approach was motivated by empirical evidence proving that models driven by a Brownian motion have several shortcomings, such as inability to incorporate jumps and leptokurticity into the price dynamics. Non-homogeneous Lévy processes and the change of measure techniques necessary for simplification and derivation of pricing formulae were also investigated. For robustness in numerical valuation, several transform methods were investigated and compared in terms of speed and accuracy. The models were calibrated to liquid South African data (ATM) interest rate caps using two methods of optimisation, namely the simulated annealing and secant-Levenberg–Marquardt methods.

Two numerical valuation approaches had been implemented in this study, the COS method and the fractional fast Fourier transform (FrFT), and were compared to the existing methods in the context. Our numerical results showed that these two methods are quite efficient and very competitive. We chose the COS method for calibration due to its rapidly speed and we have suggested a suitable approach for truncating the integration range to address the problems it has with short-maturity options. Our calibration results provided a nearly perfect fit, such that it was difficult to decide which model has a better fit to the current market state. Finally, all the implementations were done in MATLAB and the codes included in appendices.

Stellenbosch University https://scholar.sun.ac.za

Opsomming

Die standaard-Heath–Jarrow–Morton-raamwerk (kortom die HJM-raamwerk) is daarvoor bekend dat dit op die prysbepaling en verskansing van afgeleide finansiële instrumente vir rentekoerse toegepas kan word. Hierdie studie het die uitgebreide HJM-raamwerk geïmplementeer wat deur Eberlein en Raible (1999) bekendgestel is en waarin 'n Brown-beweging deur 'n breë klas prosesse met spronge vervang word. In die besonder is die HJM wat deur veralgemeende hiperboliese prosesse gedryf word ondersoek. Hierdie benadering is gemotiveer deur empiriese bewyse dat modelle wat deur 'n Brown-beweging gedryf word verskeie tekortkominge het, soos die onvermoë om spronge en leptokurtose in prysdinamika te inkorporeer. Nie-homogene Lévy-prosesse en die maatveranderingstegnieke wat vir die vereenvoudiging en afleiding van prysbepalingsformules nodig is, is ook ondersoek. Vir robuustheid in numeriese waardasie is verskeie transformmetodes ondersoek en ten opsigte van spoed en akkuraatheid vergelyk. Die modelle is vir likiede Suid-Afrikaanse data vir boperke van rentekoerse sonder intrinsieke waarde gekalibreer deur twee optimiseringsmetodes te gebruik, naamlik die gesimuleerde uitgloeimetode en die sekans-Levenberg–Marquardt-metode.

Twee benaderings tot numeriese waardasie is in hierdie studie gebruik, naamlik die kosinus-metode en die fraksionele vinnige Fourier-transform, en met bestaande metodes in die konteks vergelyk. Die numeriese resultate het getoon dat hierdie twee metodes redelik doeltreffend en uiters mededingend is. Ons het op grond van die motiveringspoed van die kosinus-metode daardie metode vir kalibrering gekies en 'n geskikte benadering tot die trunkering van die integrasiereeks voorgestel ten einde die probleem ten opsigte van opsies met kort uitkeringstermyne op te los. Die kalibreringsresultate het 'n byna perfekte passing gelewer, sodat dit moeilik was om te besluit watter model die huidige marksituasie die beste pas. Ten slotte is alle implementerings in MATLAB gedoen en die kodes in bylaes ingesluit.

Stellenbosch University https://scholar.sun.ac.za

Acknowledgements

First and foremost I would like to thank my heavenly Father for giving me strength, courage and a beautiful mind. I know without His divine favour I cannot achieve anything.

I sincerely thank my supervisor, Dr Peter Ouwehand, for his patience, support, comments, suggestions, enthusiasm and immense knowledge. Numerous conversations with him had added much value in my academic life.

I would like to extend my profound gratitude to Prof. Dr. Ernst Eberlein and Prof. Ronnie Becker for various insightful comments. I also appreciate Prof. Eberlein for sending me his articles.

I thank my family for the trust and confidence they have placed in my life. To my late great-grandmother, Paulina Kalomho: there is no mathematical expression yet to show my appreciation to you; you are still fresh in my mind.

To my mentor, Mr Veston Malango, I acknowledge all your directions. Thank you for trusting me and having faith in me. You are an amazing grace to me.

Lastly, I acknowledge the sponsorship from the Namibia Financial Institutions Supervisory Authority (NAMFISA) and Stellenbosch University's merit award from its International Office.

Stellenbosch University https://scholar.sun.ac.za

Dedication

To my mentor,
Mr Veston Malango.

Stellenbosch University https://scholar.sun.ac.za

Contents

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

List of Figures

Stellenbosch University https://scholar.sun.ac.za

List of Tables

Stellenbosch University https://scholar.sun.ac.za

Chapter 1

Introduction

1.1 Motivation

The interest rate derivatives market has expanded widely in the last few years due to advances in technology. Currently, the over-the-counter (OTC) market is dominated by interest rate products. In 2013, the Bank for International Settlements (BIS) showed that, out of US$20 158 billions derivative contracts over US$15 683 billion's worth of interest rate derivatives were outstanding in the global derivative market. This corresponds to $77,8\%$ of outstanding derivative contracts. As Table 1.1 below indicates, a large number of interest rate contracts are outstanding in the global OTC market. This demands an appropriate, sophisticated model for interest rates, more especially for the management of risk that results from the use of interest rate models.

Table 1.1: Amount of outstanding derivatives globally

Derivatives	**Notional (in billions)**	**In % of total**
Foreign exchange contracts	2 613	13
Equity-linked contracts	707	3, 5
Interest rate contracts	15 683	77, 8
Commodity contracts	394	2
Credit derivatives	732	3, 6
Other derivatives	29	0, 14

Source: BIS (2013)

The OTC interest rate derivatives consist primarily of forward rate agreements, interest rate swaps, caps and floors. Pricing of bond options is a significant problem in Financial Mathemat-

Stellenbosch University https://scholar.sun.ac.za

ics research since caplets and floorlets can be expressed as options on zero-coupon bonds while swaptions can be expressed as options on coupon bearing bonds. The classical Heath–Jarrow–Morton (HJM) framework for no-arbitrage pricing is driven by a Brownian motion. However, nowadays this model may be inadequate due to its incapability to capture newly observed market features. Market interest rates also present some features which are not consistent with a Brownian motion. Strictly speaking, a Brownian motion which is based on a normal distribution fails to fit the observed return distribution of zero-coupon bonds (see Eberlein and Prause; Eberlein and Raible, 1999; Raible, 2000). It is likely that applying an inappropriate process as a driving influence for a model may result in mispricing and model misspecification. While mispricing happens if the market prices are different from the predicted model price, model misspecification occurs if one chooses to model derivatives and hedge with an inappropriate model.

Models for interest rate derivatives have a long and rich history. A good model for interest rates should have the following general features:

(a) robustness;

(b) accurate valuation of market instruments;

(c) ease of calibration to the market data.

1.2 Background study

Traditionally, the term structure of interest rates was modelled using short rate models. Short rate models were classified into two groups, namely equilibrium models and no-arbitrage models. One of the earliest interest rate models used in the fixed income market is a type of Ornstein–Ulhenbeck process for short rates that was pioneered by Vasiček (1977) and is driven by a Brownian motion as a source of randomness. It has short rate dynamics given by

$$dr_t = a(b - r_t)dt + \sigma dW_t, \quad a, b, \sigma \in \mathbb{R}^+, \tag{1.2.1}$$

where a is the speed parameter of mean reversion to b, b is the long run average of the short rate and σ is a diffusion coefficient.

This process r_t is a mean reversion process because if $r_t < b$ then r_t increases whereas if $r_t > b$, r_t decreases, thereby pulling back to the mean level. Although this model is analytically tractable, one of its main concerns is that interest rates can become negative with positive probability, which is not a desirable feature for any interest rate model. Some modifications to Equation 1.2.1 have been implemented. These include the idea of Cox *et al.* (1985) to include

Stellenbosch University https://scholar.sun.ac.za

a square root term in the diffusion coefficient of the Vasiček short rate process to make sure that rates do not become negative, i.e.

$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dW_t, \tag{1.2.2}$$

provided that the Feller condition

$$2ab > \sigma^2$$

holds.

The long run rate b in Equations 1.2.1 and 1.2.2 can be made to be time dependent. This makes it possible for the short rate process to fit the current term structure at any time t. Consequently, for time dependent mean reversion level b, the resulting process for Equations 1.2.1 and 1.2.2 are called the Hull–White extended Vasiček (for $\gamma = 0$ in Equation 1.2.3) and the extended Cox–Ingersoll and Ross (CIR) (for $\gamma = \frac{1}{2}$ in Equation 1.2.3) respectively. Hence we have

$$dr_t = a(b_t - r_t)dt + \sigma r_t^{\gamma}dW_t. \tag{1.2.3}$$

Unfortunately, despite the fact that the above short rate models fit the initial term structure well, the literature shows that in general the market prices of interest rate derivatives are inconsistent with short rate models. This is mainly because there is no guarantee that these short rate models will continue to provide sensible prices and volatilities as they evolve. In other words, short rate models cannot properly capture the sophisticated movements of interest rate curves. Ultimately, this makes it difficult to price and hedge interest rate products whose values depend on the shape of the yield curve. This created a necessity for an alternative model for the entire yield curve. Heath *et al.* (1992) postulated a flexible framework for describing the change in the whole yield curve in terms of instantaneous forward rates (see also Section 2.2), given by

$$df(t,T) = \alpha(t,T,\omega)dt + \sigma(t,T,\omega)dW_t,\ t < T, \tag{1.2.4}$$

where W_t is a d-dimensional Brownian motion, the drift $\alpha(t,T,\omega)$ and volatility $\sigma(t,T,\omega)$ which may depend on sample paths ω.

This is the well-known Heath–Jarrow–Morton (HJM) framework in which Brownian motion is the source of uncertainties. Its application is quite popular in both academia and industry as it incorporates many of the short rate models discussed above. As mentioned above, a model driven by a Brownian motion is incapable of capturing some stylised features from the fixed income market. These features include the inabilities to reproduce the volatility smiles observed in the market, excess kurtosis and fatter tail distributions and jumps in interest rates.

Stellenbosch University https://scholar.sun.ac.za

Several modifications to the HJM framework have been made to construct a model which incorporates as many of these stylised features as possible. The first implementation was a model that can incorporate jumps given by the factor model,

$$dr_t = a(b - r_t)dt + \sigma dW_t + J_t dN_t, \tag{1.2.5}$$

where N_t is a Poisson process and J_t is a random jump size at time t, (see Das, 1998). Similar extensions were carried out by Shirakawa (1991) and Björk (1995), who added a jump component to the dynamics of zero-coupon bonds. However, this also increased computational complexity as there is no a uniform way of specifying J_t, such that we are still within the HJM framework.

The origin of inclusion of jumps in the HJM model is the work of Björk *et al.* (1997) and was later extended by Eberlein and Raible (1999), who replaced the Brownian motion (W_t) in the explicit bond price formula by a Lévy process L_t. Since forward rates can be obtained from the bond prices, and short rates from the forward rates, under a Lévy process L_t with non-negative increments, the short rate r_t in Equation 1.2.3 is non-negative and given by

$$dr_t = a(b_t - r_t)dt + \sigma dL_t. \tag{1.2.6}$$

One of the most interesting facts about this approach is that it is mathematically tractable. It is also proven that under certain conditions this approach satisfies no-arbitrage conditions. The main advantage of using Lévy processes to model the term structure of interest rates is that it gives more realistic picture of price movements on the level of the micro-structure, (see Eberlein and Raible, 1999, p. 1). In the fixed income market, Lévy processes generalised Brownian motion and as a result generalised the HJM framework.

1.3 Problem statement

Term structure of interest rates is an extremely important element of Finance. It is the information contained in the forward rates, short rates and yield curve observed from the market; a measure of how different rates with different maturities are related. It is one of the most important indicators for pricing contingent claims, determining the cost of capital and managing financial risk.

Most of this work is inspired by the research works of Eberlein and Raible (1999), Raible (2000) and Kluge (2005) and investigation into HJM models driven by a generalised hyperbolic (GH) motion and by a Brownian motion. These investigations include:

(a) model calibration to the market data;

(b) pricing methods;

Stellenbosch University https://scholar.sun.ac.za

(c) hedging analysis.

Unlike stock markets, where modelling of financial securities is restricted to a finite number of traded assets, the market of bonds consists of the whole term structure of interest rates, which (theoretically) is an infinite dimensional object, i.e. a continuum of financial securities. For this reason, bond markets demand rigorous mathematics due to the stochastic dependence structure between these securities. As a result, extensive research in this field has been developed. As for the stock market, empirical studies showed that Brownian motion fails to describe the evolution of zero-coupon bonds.

The purpose of this study is firstly to present the theories of the Lévy term structure of interest rates pioneered by Eberlein and Raible (1999) and general extension study by (Eberlein *et al.*, 2005; Kluge, 2005). Secondly, we investigate the robustness of numerical valuation methods, including the application of the cosine and fractional FFT methods, and compare these with algorithms developed by Raible (2000). Once the better pricing method is identified, we calibrate the model to South African (ATM) interest rate caps/floors and swaptions. Finally, we give a statement on hedging in Lévy bond market.

The most difficult question in term structure modelling is which models are good enough. The answer to this question varies depending on the particular purpose and application for that model. In general, if a model is to be used for pricing derivative products then calibration and hedging play a major role. Calibration is the method of estimating model unobserved parameters by making sure that the "distance" between the model prices and market observed prices is as small as possible. Hedging, on the other hand, is a method to minimise the probability of loss from a particular contingent claim by trading the underlying hedging instruments. It is an important practice in modern trading and risk management. Model risk is defined as the risk that arises from applying an inadequate model, i.e. a risk that arises when a "wrong" model is used to price and hedge a derivative. Applying the "correct" model in both pricing and risk management is a desirable and fundamental concept. But how do we determine the correct model? Does it involve calibration alone? We are aware that different methods of calibration may result in different sets of parameters. Wrong calibration is one of the main reasons for model risk. What are the criteria for deciding upon the best model? What about hedging? Our desire is to have a model that we can easily calibrate to fit the market quoted data and to easily set up an appropriate hedging method.

1.4 Literature review

The theory of HJM driven by a Brownian motion is well studied and implemented in both academia and practice. The assumptions that interest rates follow a pure diffusion process are

Stellenbosch University https://scholar.sun.ac.za

doubted nowadays due to observations of jumps or spikes in the bond market. It is shown that the empirical distribution of interest rates exhibits excess kurtosis, skewness and higher moments which are inconsistent with a normal distribution, (see Eberlein, 2001; Eberlein and Raible, 1999; Raible, 2000, for empirical motivation). To be able to capture all observed features in interest rates, Shirakawa (1991) conducted a study who included a pure jump component in the forward rate dynamics to allow for the occurrence of jumps. His main ideas were to consider a model driven by a standard Brownian motion and to add a Poisson process with constant jump intensity. This process is known as the jump-augmented HJM model. It was also shown that a model with constant jump intensities is not realistic enough. Further modification of the jump-augmented HJM was done by Jarrow and Madan (1995), who considered all jump intensities to be path-dependent. Björk *et al.* (1997) extended the HJM model by consider forward rate dynamics driven by a Brownian motion and random measures (general semi-martingale) with finite compensator.

This study focuses on a more general class of HJM models which have general Lévy process as driving processes (henceforth called Lévy HJM). Lévy processes are very general stochastic processes with stationary independent increments that can incorporate jumps, fatter tail and high peak distribution. The theory on Lévy HJM was introduced by Eberlein and Raible (1999) and Raible (2000) as a general extension of the HJM jump-diffusion model of Björk *et al.* (1997). The key idea of Lévy HJM is to replace the Brownian motion in the HJM models with a general Lévy process under some conditions. As mentioned by Eberlein and Raible (1999) however, we do not replace the Brownian motion in the stochastic differential equation of bond dynamics, but in the explicit bond price formula. The reason for this is that replacing a Brownian motion by a Lévy process in the differential equation will lead to a Doléans-Dade exponential solution (see Theorem 3.7.2), which tends to produce negative prices for Lévy processes with negative jumps greater than one.

Lévy HJM models have become an important subject in Mathematical Finance literature. Eberlein *et al.* (2005) and Kluge (2005) extended the Lévy HJM framework by applying a more general class of driving processes known as non-homogeneous or time-inhomogeneous Lévy processes. These are processes with independent increments and absolutely continuous characteristics (PIIAC). Kluge (2005) suggested rather than considering a general Lévy process, using a stochastic integral of a deterministic function with respect to a non-homogeneous Lévy process, known as the driving process. In other words, the use of non-homogeneous Lévy process is due the fact that the change of measure is vital for simplicity and for derivation of pricing formulae. The reason for this is because driving processes of non-homogeneous Lévy process are invariant under the change of measure; the change of measure is not structure preserving for homogeneous Lévy processes, (see Eberlein *et al.*, 2005).

Stellenbosch University https://scholar.sun.ac.za

Furthermore, Kluge (2005) calibrated models to the market implied volatilities for ATM caps and swaptions, and considered two Lévy-driven models, one driven by a homogeneous and the other by a non-homogeneous Lévy process. His findings were that a model driven by a non-homogeneous Lévy process provides a better fit to implied volatility.

Although models driven by Lévy processes are well studied and tractable, they often lack closed-form solution to option valuations. This creates a trade-off between numerical and analytical option value evaluations.

One of the crucial fundamental problems in Financial Mathematics is the explicit computation of derivative prices. There is a greater computational complexity for option valuations. Efficient numerical methods are required to compute derivative prices accurately and efficiently for model calibration to liquid market data for interest rates caps/floors and swaptions. Usually an ideal tool for numerical computation in finance is the Monte Carlo method, which is capable of derivative valuation of any kind. This numerical method has drawback its lack of computational speed.

Raible (2000) developed algorithms for any general contingent claims valuation based on bilateral Laplace transforms. He considered the convolution of an arbitrary pay-off function and density function and found the bilateral Laplace transform of the product, which can then be computed using the fast Fourier transform (see Carr and Madan, 1999). Since then, numerous papers have tried to present alternative methods to improve the computational complexity. Examples of these are a direct modification of Raible's ideas by Eberlein and Kluge (2006), who derived pricing formulae for caps, floors and swaptions using ideas of a convolution representation; Kuan and Webber (2001), who proposed the use of random trinomial lattice; and some Fourier based methods for option valuation under the Lévy HJM framework, discussed in Eberlein *et al.* (2010). In this study we apply the following methods of pricing:

(a) Monte Carlo method (see Glasserman, 2003).

(b) Fourier-based methods (see Eberlein *et al.*, 2010).

(c) FFT method (see Raible, 2000).

(d) FrFT method (see Chourdakis, 2004).

(e) COS method (see Fang and Oosterlee, 2008).

A financial model which is used for pricing cannot be separated from hedging. The knowledge of pricing a derivative contract must be accompanied by the knowledge of preventing the associated risk. Generally, Lévy HJM presents an exponential Lévy market which is incomplete.

Stellenbosch University https://scholar.sun.ac.za

Eberlein *et al.* (2005) showed that in the one-dimensional case, the Lévy HJM framework is complete, hence perfect hedging is possible. The original study of Björk *et al.* (1997) showed that in models with processes that exhibit jumps, interest rates cannot be perfectly hedged using zero-coupon bonds. This means there are some risk factors that affect the price for derivative product but do not affect the underlying variable. This is termed unspanned stochastic volatility (USV). This implicitly means, except for a one-dimensional case, generally, Lévy HJM will introduce unspanned stochastic volatility, which makes hedging nearly impossible.

Moreover, although it is quite popular that in one-dimensional Lévy HJM framework there is a unique risk-neutral measure, little is known about hedging interest rate products within this framework. To the best of our knowledge, only one paper (see Vandaele, 2010, Chapter 8) discussed hedging in this market.

1.5 Thesis structure

The rest of this study is structured as follows: Chapter 2 introduces the fixed income derivatives, trading methodologies, the notion of no-arbitrage, risk-neutral valuation principles and the financial instruments that will be required in the subsequent chapters. Also in Chapter 2, we introduce interest rate derivatives. We show that a cap/floor is a portfolio of options on zero-coupon bonds while a swaption is an option on a portfolio of coupon bonds. We derive pricing formulae for these interest rate derivatives.

Chapter 3 outlines the properties that define Lévy processes and discusses the generalised hyperbolic processes. It introduces an example of a non-homogeneous Lévy process as represented by a stochastic integral of a deterministic function with respect to a homogeneous Lévy process. Some basic theorems and results in Itô calculus theory (such as exponential semimartingale and the Girsanov Theorem for Lévy processes) will be introduced, as these play a crucial part in the change of pricing measure.

Chapter 4 discusses the Lévy HJM framework. It begins with the introduction to the classical theory of HJM models and the general extension known as the driving process. This includes the description of the driving process and volatility structure. We shall show that discounted bond prices process are martingales and present an equivalent HJM drift condition for Lévy HJM models. The change of numéraire techniques are introduced and we show how to compute the risk-neutral expectation without need of change of numéraire using the Monte Carlo method.

In Chapter 5, various numerical valuation methods based on Fourier methods are implemented and compared in terms of speed and accuracy . In this chapter we apply two methods to the literature Lévy term structure modeling, namely the COS method and fractional Fourier

Stellenbosch University https://scholar.sun.ac.za

transform. Our unique approach is that we have derived our pricing formula from a well-known Parseval's Theorem in probability theory.

Chapter 6 presents model calibration using two methods of optimisation, namely the secant-Levenberg–Marquardt method and the Simulated Annealing method. It begins with the extension of derivation of the pricing formulae for interest rates derivatives introduced in Chapter 2.

Chapter 7 discusses the lack of hedging literature in Lévy HJM and limitations.

Chapter 8 gives an overview, concluding remarks, and direction for future research. Finally, in Appendices, we present some theory and MATLAB codes.

Stellenbosch University https://scholar.sun.ac.za

Chapter 2

Introduction to fixed income derivatives

Fixed income market are types of markets whereby market participants trade contracts in which cash-flows are prescribed in future time. The fundamental problem is to know how to value these contracts and prevent the associated risks.

This chapter introduces the basic concepts of stochastic modelling in the theory of interest rates, no-arbitrage options valuation and the interest rate derivatives. We shall begin by reminding the reader on some essential definitions, a discussion on the trading strategies used in the fixed income market of interest rates, the notion of numéraire and risk-neutral valuation principle.

For basic theory on interest rate models, we refer the reader to the works of Musiela and Rutkowski (2004), James and Webber (2000), Bingham and Kiesel (2004), Brigo and Mercurio (2006) and Björk (2009).

2.1 Definitions

Modelling any financial asset requires risk-neutral pricing, whereby the price of a security is obtained by taking the expectation of its discounted pay-off under a risk-neutral measure. The central concept in risk-neutral valuation is the absence of arbitrage opportunities. Recall that an arbitrage opportunity is a self-financing strategy with an initial value of zero, which almost surely produces a non-negative final value and has a strictly positive probability of positive final value.

Throughout this study, we let $\mathbb{T} \in \mathbb{R}$ be a fixed maximum finite time horizon for all market activities and $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ represent a filtered probability space defining the framework that gives the characterization of uncertainty in the economy. Here Ω represents the state or sample space, $\mathcal{F}$ is the σ-algebra of all events, $\mathbb{P}$ is the real-world probability measure and $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,\mathbb{T}]}$ is the filtration, assumed to satisfy the usual conditions over the interval $[0, \mathbb{T}]$.

Stellenbosch University https://scholar.sun.ac.za

We let the bank account process be the numéraire, i.e.

$$B_t = \exp\left\{\int_0^t r_s\, ds\right\}, \quad \forall t \in [0, \mathbb{T}], \tag{2.1.1}$$

where r_t is the short rate.

In our set-up, there are infinitely many zero-coupon bonds traded in the market, and hence the below definition for an equivalent martingale measure (EMM) applies to a continuum of financial securities (unlike in the case of equity, where it only applies to a finite collection of stocks).

Definition 2.1.1 *A probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ is an equivalent martingale measure (EMM) if $\mathbb{Q} \sim \mathbb{P}$ and discounted bond price process Z_t is martingale, i.e.*

$$Z_t = \frac{P(t,T)}{B_t} = \mathbb{E}_{\mathbb{Q}}\left[\frac{P(T,\cdot)}{B_T}\middle|\mathcal{F}_t\right], \quad \forall 0 \le t \le T \le \mathbb{T}.$$

Definition 2.1.2 *A probability measure $\mathbb{Q}^T$ on $(\Omega, \mathcal{F})$ equivalent to $\mathbb{Q}$ with the Radon-Nikodym derivative given by*

$$\frac{d\mathbb{Q}^T}{d\mathbb{Q}} = \frac{B_T^{-1}}{\mathbb{E}_{\mathbb{Q}}[B_T^{-1}]} = \frac{1}{B_T P(0,T)}, \quad \mathbb{Q} - a.s.,$$

is called the forward martingale measure for the settlement date T.

The above Radon-Nikodym derivative when restricted to a σ-field $\mathcal{F}$ for every $t \in [0, T]$ gives a density

$$R_t := \frac{d\mathbb{Q}^T}{d\mathbb{Q}} = \mathbb{E}_{\mathbb{Q}}\left[\frac{B_0 P(T,T)}{B_T P(0,T)}\middle|\mathcal{F}_t\right] = \frac{B_0 P(t,T)}{B_t P(0,T)}.$$

Definition 2.1.3 *A security market is arbitrage-free if there are no arbitrage opportunities.*

Definition 2.1.4 *A contingent claim X is a financial instrument which at maturity T pays an amount with pay-off function $\Phi(X_T)$, where X_T is an $\mathcal{F}_T$-measurable random variable which is bounded below.*

Definition 2.1.5 (European put option) *Let S_T be the price of a financial instrument at maturity time T. A put option is a contingent claim with a strike K whose pay-off at maturity date T is given by*

$$\Phi_{put}(S_T) = \max\{K - S_T, 0\}. \tag{2.1.2}$$

The pay-off for a call option on S_T is given by

$$\Phi_{\text{call}}(S_T) = \max\{S_T - K, 0\}. \tag{2.1.3}$$

Stellenbosch University https://scholar.sun.ac.za

Definition 2.1.6 *The arbitrage price process of any contingent claim X is given by the risk-neutral valuation formula, i.e., the value of a contingent claim at time t is given by*

$$X_t = B_t \mathbb{E}_{\mathbb{Q}} \left[\frac{\Phi(X_T)}{B_T} \middle| \mathcal{F}_t \right] \tag{2.1.4}$$

A major problem encountered with risk-neutral valuation is that the risk-neutral measure is not unique in most cases which causes prices to differ. For this reason, we ought to work in a complete market.

Definition 2.1.7 *An arbitrage-free market is complete if and only if there exists a unique equivalent martingale measure under which discounted asset prices are martingales.*

The above definition is rephrased from a well-known fundamental theorem of asset pricing.

2.2 Financial instruments

In this section we introduce the no-arbitrage trading strategies to derive formulae for basic financial products.

Definition 2.2.1 *Fix a maturity $T < \mathbb{T}$. A zero-coupon bond with maturity T (T-bond) is a financial instrument paying 1 unit of currency to the holder at time T. Its value at time $t \leq T$ is denoted by $P(t,T)$. The process*

$$t \mapsto P(t,T) \quad t \leq \mathbb{T}$$

is adapted to the filtration $\mathbb{F}$ with $P(T,T) = 1$.

We assume that there are zero-coupon bonds for all maturities $T \in [0, \mathbb{T}]$. Since the main task is to find the arbitrage-free prices of interest rate derivatives, such as bond options, caps/floors and swaptions, we firstly look at the following no-arbitrage trading strategy. Let $t < T < U$ be a fixed time interval.

Table 2.1: No-arbitrage arguments

Strategy	**Quantity at time t**	**T**	**U**
short	$1 \ \ T-$bond	pay 1.00	–
long	$\frac{P(t,T)}{P(t,U)}\ U-$bonds	–	receive $\frac{P(t,T)}{P(t,U)}$
Net income	0.00	-1.00	$\frac{P(t,T)}{P(t,U)}$

Stellenbosch University https://scholar.sun.ac.za

- At time t, sell a bond maturing at time T, and use the cash to buy $\frac{P(t,T)}{P(t,U)}$ bonds maturing at time U. The net investment is zero.
- At time T, pay 1 unit to redeem T-bond.
- At time U, collect $\frac{P(t,T)}{P(t,U)} \times 1$ units from U-bond.
- This implies 1 unit deposited at time T leads to a payment of $\frac{P(t,T)}{P(t,U)}$ at time U.

Definition 2.2.2 *The interest rate that can be locked in today (at time t) for a future period $[T, U]$ is known as the forward rate and it is denoted by $R(t, T, U)$.*

According to the above definition, forward rates are contracted rates at time t which become available at time T and end at time U.

From the trading arguments in Table 2.1, it looks as if a deposit of 1 unit is made at time T and earns interest R for the period $[T, U]$, i.e. to avoid arbitrage opportunities

$$1.e^{R(U-T)} = \frac{P(t,T)}{P(t,U)}$$

$$\Longrightarrow R(t,T,U) = -\frac{\log P(t,U) - \log P(t,T)}{U-T}.$$

Definition 2.2.3 *The forward rate that can be locked in at time t for an infinitesimal interval $[T, T + dT]$, given by*

$$f(t,T) = \lim_{\Delta t \to 0} R(t,T,T+\Delta T) = -\frac{\partial}{\partial T} \log P(t,T)$$

is called the instantaneous forward rate.

The forward rate $R(t, T, U)$ is assumed here to be continuously compounded. The simply compounded versions of forward rates are called the forward LIBOR rates (London Interbank Offered Rate) and are denoted by $L(t, T, U)$. For time $t \leq T \leq U$, the forward LIBOR rate is defined by

$$L(t,T,U) = \frac{1}{U-T}\left(\frac{P(t,T)}{P(t,U)} - 1\right). \tag{2.2.1}$$

We define a function $T \mapsto f(t, T)$ for fixed t to be the forward rate curve at time t. Since initial forward rates are directly observed from the market, one can recover bonds from the forward rate curve using the relation

$$P(t,T) = \exp\left(-\int_t^T f(t,s)\, ds\right). \tag{2.2.2}$$

Stellenbosch University https://scholar.sun.ac.za

The short rate process, defined by $r_t = r(t) = f(t,t)$, is the instantaneous short rate at time t.

The forward bond price process at time t for a bond to be bought at time T and maturing at time U is an amount $\mathcal{B}(t,T,U)$ contracted for at time t, to be paid at time T, for a bond with maturity time U. It is given by no-arbitrage arguments

$$\mathcal{B}(t,T,U) = \frac{P(t,U)}{P(t,T)}, \quad \text{for } 0 < t < T < U < \mathbb{T}. \tag{2.2.3}$$

2.3 Interest rate derivatives

The objective of this subsection is to define interest rate derivatives and their cash-flows. These include interest rate caps/floors and swaptions which form a basic and useful tool in managing the risk of any financial institution. As we have already seen in Chapter 1, the main traded and outstanding fixed income instruments in the global financial market include bonds and swaps. Essentially, caps/floors and swaptions are options on these instruments.

Interest rate caps/floors have a simple relation with zero-coupon bonds, while swaptions are associated with coupon paying bonds. In other words, caps/floors and swaptions can be modelled depending on a single underlying variable. Ultimately, this makes the valuation methods relatively easy due to the availability of pricing formulae for bonds in various interest rate models which are obtained via no-arbitrage principles. We follow the works of Musiela and Rutkowski (2004) and Brigo and Mercurio (2006).

2.3.1 Options on forward rates

Definition 2.3.1 *A caplet is a call option on LIBOR rate with strike rate κ and maturity time U. Its pay-off is given by*

$$\Phi_{\mathit{Caplet}}(L(t,T,U)) = \max\{L(t,T,U) - \kappa, 0\}. \tag{2.3.1}$$

Similarly, a floorlet is a put option on a LIBOR rate. Its pay-off is given by

$$\Phi_{\text{Floorlet}}(L(t,T,U)) = \max\{\kappa - L(t,T,U), 0\}. \tag{2.3.2}$$

A forward start cap/floor is a strip of small options called caplets/floorlets. A caplet contracted on time period $[T_{i-1}, T_i]$, $\delta_i = T_i - T_{i-1}$ settled in arrears pays the holder an amount of $\delta_i(L(T_{i-1}, T_{i-1}, T_i) - \kappa)^+$ at time T_i, where L is the simple forward LIBOR rate determined at time T_{i-1} and given by Equation 2.2.1. Similarly, a floorlet pays the holder an amount of $\delta_i(\kappa - L(T_{i-1}, T_{i-1}, T_i))^+$.

Stellenbosch University https://scholar.sun.ac.za

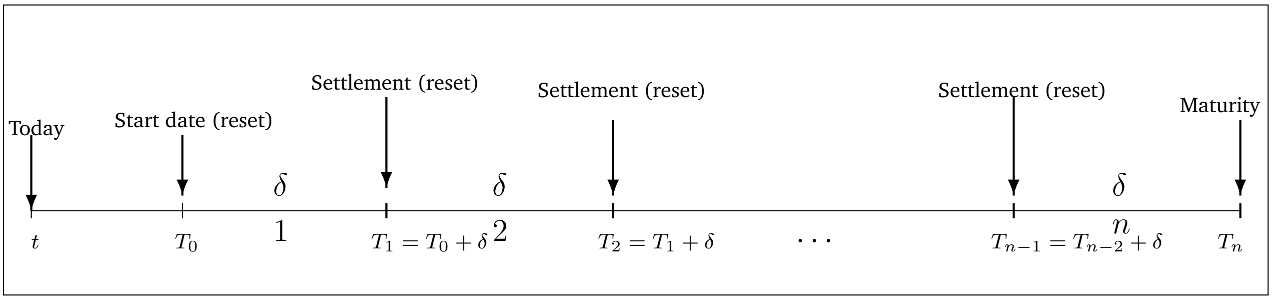
A cap/floor protects the holder of the contract against rising(falling) LIBOR rates because it makes sure that interest is to be paid if the floating rate L_t exceeds(remain below) a cap/floor rate κ for all $t < \mathbb{T}$.

A floorlet can be statically replicated using a caplet and a Forward Rate agreement, and for this reason, it is usually suffices to focus on caplet pricing, i.e. a floor can be obtained from cap-floor parity relation:

cap(t) − floor(t) = FRA.

Suppose $\mathcal{T} = \{T_0 < T_1 < \cdots < T_n\}$ is a sequence of payment dates, frequently called the tenor structure for a cap.

Tenor structure



We grouped these into two sets of dates, namely reset dates $(T_0, T_1, \cdots, T_{n-1})$ and payment dates $(T_1, T_2, \cdots, T_n)$. The arbitrage-free value of a caplet on $[T_{i-1}, T_i]$ is given by

$$\begin{aligned}
\mathbb{V}_{\text{caplet}}(t, T_{i-1}, T_i) &= \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^{T_i} r_s\, ds}\delta_i(L(T_{i-1}, T_{i-1}, T_i) - \kappa)^+\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^{T_{i-1}} r_s\, ds}e^{-\int_{T_{i-1}}^{T_i} r_s\, ds}\delta_i(L(T_{i-1}, T_{i-1}, T_i) - \kappa)^+\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^{T_{i-1}} r_s\, ds}\mathbb{E}_{\mathbb{Q}^{T_{i-1}}}\left[e^{-\int_{T_{i-1}}^{T_i} r_s\, ds}\right]\delta_i(L(T_{i-1}, T_{i-1}, T_i) - \kappa)^+\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^{T_{i-1}} r_s\, ds}P(T_{i-1}, T_i)\delta_i\left(\frac{1}{\delta_i}\left(\frac{1}{P(T_{i-1}, T_i)} - 1\right) - \kappa\right)^+\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^{T_{i-1}} r_s\, ds}P(T_{i-1}, T_i)\left[\left(\frac{1}{P(T_{i-1}, T_i)} - 1\right) - \delta_i\kappa\right]^+\right] \qquad (2.3.3) \\
&= \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^{T_{i-1}} r_s\, ds}\left[1 - P(T_{i-1}, T_i) - \delta_i\kappa P(T_{i-1}, T_i)\right]^+\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^{T_{i-1}} r_s\, ds}\left[1 - P(T_{i-1}, T_i)(1 + \delta_i\kappa)\right]^+\right] \\
&= (1 + \delta_i\kappa)\mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^{T_{i-1}} r_s\, ds}\left(\frac{1}{1 + \delta_i\kappa} - P(T_{i-1}, T_i)\right)^+\right] \\
&= (1 + \delta_i\kappa)P(t, T_{i-1})\mathbb{E}_{\mathbb{Q}^{T_{i-1}}}\left[\left(\frac{1}{1 + \delta_i\kappa} - P(T_{i-1}, T_i)\right)^+\right].
\end{aligned}$$

Stellenbosch University https://scholar.sun.ac.za

In the last equality we have used the change of measure in Equation 2.1.2.

Therefore to value an i^{th} caplet simply evaluate $1 + \delta_i\kappa$ put options with strike price $\frac{1}{1+\delta_i\kappa}$ and maturity time T_{i-1} on a zero-coupon bond maturing at time T_i. Similarly, a i^{th}-floorlet is equivalent to $1 + \delta_i\kappa$ call options with maturity T_{i-1} and strike price $\frac{1}{1+\delta_i\kappa}$ on a zero-coupon bond maturing at time T_i.

2.3.2 Options on interest rate swap

Definition 2.3.2 *A swap is a sequence of n interest rates which consist of starting and ending dates $0 \leq T_1 < T_2 < \cdots < T_{n+1} \leq \mathbb{T}$, $\delta_i = T_i - T_{i-1}$ for all $1 \leq i \leq n$. At the exchange date, the payer gets the interest rate payment $\delta_i(L(T_{i-1}, T_{i-1}, T_i) - \kappa)$ at time T_i and the receiver gets $\delta_i(K - L(T_{i-1}, T_{i-1}, T_i))$ at time T_i.*

Interest rate swap is a contract intended for exchanging interest rates. One side of cash-flows pays a fixed amount (known as a fixed leg) while the other side pays a floating LIBOR rate (known as a floating leg), which is determined in advance. One side is called a payer swap if it pays a fixed and receives a floating amount; the other side is called a receiver swap if it receives a fixed amount and pays a floating amount.

A swap contract is specified by the reset dates, payment dates and fixed rate of the contract. The fixed payment of $\delta_i\kappa$ is settled at payment dates. The floating payment of $\delta_i L(T_{i-1}, T_{i-1}, T_i)$ is also settled at payment dates but is determined at the previous reset date.

Consider the interval $[T_{i-1}, T_i]$. Then at time T_i, the floating leg pays

$$\delta_i L(T_{i-1}, T_{i-1}, T_i)$$

and the contract is worth

$$P(T_{i-1}, T_i)\delta_i L(T_{i-1}, T_{i-1}) = P(T_{i-1}, T_i)\delta_i \left(\frac{1}{\delta_i}\left(\frac{1}{P(T_{i-1}, T_i)} - 1\right)\right) = 1 - P(T_{i-1}, T_i).$$

For $t < T_{i-1}$, this is worth $1 - P(t, T_i)$.

Hence, the time t-value of all floating rate payments is

$$\sum_{i=1}^{n} P(t, T_{i-1}) - P(t, T_i) = P(t, T_n) - P(t, T_0)$$

whereas the time t-value of fixed rate payments is

$$\sum_{i=1}^{n} P(t, T_i)\delta_i\kappa.$$

Stellenbosch University https://scholar.sun.ac.za

The value of a payer swap therefore is equivalent to a portfolio short position a coupon bond with coupon rate κ and long position a floating rate. Hence, its time t-value is

$$\sum_{i=1}^{n} P(t,T_{i-1}) - P(t,T_i) - \sum_{i=1}^{n} P(t,T_i)\delta_i\kappa = P(t,T_0) - P(t,T_n) - \sum_{i=1}^{n} P(t,T_i)\delta_i\kappa$$

$$= P(t,T_0) - \left(P(t,T_n) + \sum_{i=1}^{n} P(t,T_i)\delta_i\kappa\right). \tag{2.3.4}$$

The terms in the bracket can be seen as a coupon bond with coupon rate κ. The forward swap rate is defined to be

$$S_t = \frac{P(t,T_0) - P(t,T_n)}{\sum_{i=1}^{n} P(t,T_i)\delta_i}. \tag{2.3.5}$$

The swap rate is that rate $\kappa = S_{T_0}$ such that a swap starting at T_0 has an initial value of zero, i.e.

$$S_{T_0} = 0 = \frac{P(T_0,T_0) - P(T_0,T_n)}{\sum_{i=1}^{n} P(T_0,T_i)\delta_i\kappa} = \frac{1 - P(T_0,T_n)}{\sum_{i=1}^{n} P(T_0,T_i)\delta_i}.$$

$$\Longleftrightarrow 1 = P(t,T_n) + \sum_{i=1}^{n} P(t,T_i)\delta_i S_{T_0}.$$

This means that a swap rate can be defined as a coupon rate for a coupon bond traded at par (called the ”par yield”).

Definition 2.3.3 *Swaptions are options on interest forward starting swaps between time T_0 and T_n. A swaption gives the holder the right but not obligation to enter into a particular swap contract (see Definition 2.3.2).*

The holder of a payer(receiver) swaption with strike rate κ and maturity T has the right to enter at time T a forward payer(receiver) swap which is settled in arrears. The maturity of the swaption usually coincides with the starting date for the interest rate swap, i.e. $T = T_0$.

We focus on payer swaptions because receiver swaptions can be found from payer-receiver swaptions parity:

$$\mathbb{V}_{\text{PS}}(t) - \mathbb{V}_{\text{RS}}(t) = \text{swap},$$

where $\mathbb{V}_{\text{PS}}(t)$ and $\mathbb{V}_{\text{RS}}(t)$ stand for the value of payer and receiver swaptions at time t respectively with the same strike rate and tenor structure.

Denote the value of a coupon bond at $t < T_0$ by

$$Z_t = \sum_{i=1}^{n} c_i P(t,T_i),$$

Stellenbosch University https://scholar.sun.ac.za

where $c_i = \kappa\delta_i$ for $i = 1, 2, \cdots, n-1$ and $c_n = 1+\kappa\delta_n$. Z_t is sometimes referred to as an annuity factor.

Denote the time t-value of the forward payer swap contract with fixed interest rate κ in Equation 2.3.4 by

$$\text{PS}(t, \kappa) = P(t, T_0) - \sum_{i=1}^{n} c_i P(t, T_i).$$

The arbitrage price of the payer swaption is

$$\begin{aligned}\mathbb{V}_{\text{PS}}(t) &= \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^T r_s\, ds}\left(P(T, T_0) - \sum_{i=1}^{n} c_i P(T, T_i)\right)^+ \middle| \mathcal{F}_t\right] \\ &= \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^T r_s\, ds}\left(1 - \sum_{i=1}^{n} c_i P(T, T_i)\right)^+ \middle| \mathcal{F}_t\right].\end{aligned}$$

Hence, a payer swaption can be seen as a put option with strike price $K = 1$ and maturity $T \leq T_0$ on the coupon paying bond (see also Musiela and Rutkowski, 2004, Chaper 13).

Hence the value of the payer swaption under the forward martingale measure $\mathbb{Q}^T$ is given by

$$\mathbb{V}_{\text{PS}}(t) = P(t, T)\mathbb{E}_{\mathbb{Q}^T}\left[\left(1 - \sum_{i=1}^{n} c_i P(T, T_i)\right)^+ \middle| \mathcal{F}_t\right]. \tag{2.3.6}$$

Stellenbosch University https://scholar.sun.ac.za

Chapter 3

Lévy processes

Most financial models are based on the assumption that asset returns follow a normal distribution. However, recent study on financial modeling provides empirical motivation that using a normal distribution to model the logarithmic returns of the underlying variable does not adequately capture some features observed in the option market. In option markets, asset prices present jumps and spikes which are not consistent with a model based on a normal distribution. In essence, a normal distribution cannot explain well the exhibited features such as the skew or smile of the implied asset return distribution of the underlying variable. Hence a model driven by a Brownian motion alone is not sufficient in modeling financial derivatives as it does not account for these non-phenomenon features and does not allow for discontinuities and jumps in the derivative price process either. To overcome this problem, the model driven by a Brownian motion is often generalised by applying processes that allow it to accurately fit the return distribution of the asset price process.

One common generalisation of Brownian motion is the use of Lévy processes, i.e., processes that have stationary, independent increments. These are a more general class of processes that are able to incorporate jumps into their dynamics. It is the jump features in Lévy processes that make these models valuable tools for financial modeling. This is because the jump component in Lévy processes is responsible for describing skew and smile for options with short time to maturity. In addition, applying Lévy processes to term structure models does not only improve the fit of the empirical distribution but also provide a better description of the interest rates movement (see Eberlein and Raible, 1999, p. 1).

This chapter introduces the mathematics behind Lévy processes. We follow closely Cont and Tankov (2004), Schoutens (2003), Eberlein (2001) and Sato (2001). Our aim is to define this class of stochastic processes and summarise the results applied in the main part of this study. We shall focus on the generalised hyperbolic model and few of its subclasses in one dimension.

Stellenbosch University https://scholar.sun.ac.za

3.1 Definitions

Let $\mathbb{T}$ be a fixed time horizon. Denote a complete stochastic basis by $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Here $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ is a filtration satisfying the usual conditions.

Definition 3.1.1 *A function $f : \Omega \mapsto \mathbb{R}$ is called a càdlàg if the limits $f(t_-) = \lim_{\Delta t\to 0} f(t - \Delta t)$ and $f(t_+) = \lim_{\Delta t\to 0} f(t + \Delta t)$ exists and $f(t) = f(t_+)$.*

From the above definition, a process $X = (X_t)_{t\geq 0}$ such that $X_0 = 0$ is said to be a càdlàg process if it is continuous on the right with limits on the left (RCLL).

Definition 3.1.2 *A real-valued stochastic process $X = (X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is called a Lévy process if the following conditions are satisfied (Sato, 2001):*

(a) X has independent increments, i.e. for any $t \geq s$, $X_t - X_s \perp\!\!\!\perp \mathcal{F}_s$.

(b) X is time homogeneous (increments are stationary), i.e. the distribution of $X_{t+h} - X_h : t \geq 0$ does not depend on h.

(c) X is stochastically continuous. This means for every $\epsilon > 0$, $\mathbb{P}(|X_{t+h} - X_h| > \epsilon) \to 0$ as $t \to 0$.

(d) The sample path $X_t(\omega)$ is right continuous with limit from the left, i.e. càdlàg almost surely.

(e) $X_0 = 0$ (almost surely).

The third condition means that sample paths are not necessary continuous and the probability of seeing jumps at any given time t is zero, i.e. discontinuities occur randomly. The fourth condition is useful in the analysis of processes with independent and stationary increments whereas the last condition is for normalisation purposes.

A stochastic process is called a Lévy process in law if it satisfies conditions $(a), (b), (c)$ and (e). It is called an additive process if conditions $(a), (c), (d)$ and (e) hold, i.e. relaxing condition (b). Furthermore, it is called an additive process in law if conditions $(a), (b)$ and (e) hold.

Every Lévy process in law has a càdlàg modification, i.e. a càdlàg process Y_t is such that $X_t = Y_t$ almost surely for every t. As a result, we restrict our discussion to Lévy processes that are càdlàg processes.

One of the simplest Lévy processes is the linear drift (-a deterministic process).

Definition 3.1.3 (Brownian motion) *A stochastic process W_t defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a $\mathbb{P}$-Brownian motion if*

Stellenbosch University https://scholar.sun.ac.za

(a) $W_0 = 0$ *almost surely*

(b) W_t *has stationary increments*

(c) W_t *has independent increments*

(d) for $0 < s < t$, $W_{t+s} - W_t \sim N(0, s)$

We should stress that an arithmetic Brownian motion as introduced in Definition 3.1.3 is the only Lévy process with continuous sample paths. In essence, Lévy processes generalise Brownian motion by relaxing the condition of continuous paths. This means that in this case increments need not be normally distributed.

Definition 3.1.4 (Convolution) *Let* μ_1 *and* μ_2 *be the probability distributions on* $\mathbb{R}^d$*. The convolution of* μ_1 *and* μ_2 *is denoted by* $\mu_1 * \mu_2$ *and it is a distribution defined by*

$$\mu(B) = \mu_1 * \mu_2(B) = \int\int_{\mathbb{R}^d} \mathbb{I}_B(x+y)\mu_1(dx)\mu_2(dy), \quad B \in \mathcal{B}(\mathbb{R}^d).$$

We denote the n*-fold convolution of* μ *by* μ^{n*}*.*

A measure on $\mathbb{R}$ induced by a random variable X is denoted by

$$\mu_X(A) = \mathbb{P}(X \in A),$$

where A is a Borel subset of $\mathbb{R}$.

Definition 3.1.5 (Characteristic function) *The characteristic function of a probability measure* μ *on* $\mathbb{R}^d$ *is defined to be a map* $\chi : \mathbb{R}^d \mapsto \mathbb{C}$ *defined by*

$$\chi(u) = \int_{\mathbb{R}^d} e^{\mathrm{i}<u,x>}\mu(dx), \quad \text{for all } u \in \mathbb{R}^d, \tag{3.1.1}$$

where $\mathrm{i} = \sqrt{-1}$ *is a complex number and* $< \cdot >$ *denotes the inner product.*

The characteristic function of a random variable X is given by

$$\chi_X(u) = \int_{\mathbb{R}} e^{\mathrm{i}ux}\mu_X(dx) = \mathbb{E}\left[e^{\mathrm{i}uX}\right].$$

The law or a distribution of a random variable X is denoted by $\mathbb{P}_X(x) = \mathbb{P}(X \leq x)$.

Proposition 3.1.6 *A distribution* μ *on* $\mathbb{R}^d$ *is said to be infinitely divisible if for every natural number* $n \geq 1$ *there exists a* n*-fold convolution of* μ_n *(Sato, 2001).*

Stellenbosch University https://scholar.sun.ac.za

The above proposition says that the law of a random variable X is infinitely divisible, if for all $n \in \mathbb{N}$ there is a random variable $X^{\left(\frac{1}{n}\right)}$ such that

$$\chi_X(u) = \left(\chi_{X^{\left(\frac{1}{n}\right)}}(u)\right)^n .$$

A random variable is said to be infinitely divisible if and only if its distribution μ is infinitely divisible.

We provide a simple example to show that the normal distribution is infinitely divisible.

Example 3.1.7 *Let $X \sim N(\mu, \sigma)$. Then*

$$\chi_X(u) = \exp\left(iu\mu - \frac{1}{2}u^2\sigma^2\right) = \left(\exp\left[iu\frac{\mu}{n} - \frac{1}{2}u^2\frac{\sigma^2}{n}\right]\right)^n = \left(\chi_{X^{\left(\frac{1}{n}\right)}}(u)\right)^n .$$

Here $X^{\left(\frac{1}{n}\right)} \sim N\left(\frac{\mu}{n}, \frac{\sigma^2}{n}\right)$.

Theorem 3.1.8 *If X_t is an additive process in law, then for every $t \geq 0$, X_t is infinitely divisible. Conversely, if ρ is an infinitely divisible distribution on $\mathbb{R}^d$, then there exists uniquely in law, a Lévy process in law X_t such that $\mathbb{P}_X = \rho$.*

Lévy processes are general stochastic processes which are fully described by their characteristic function. The following Theorem gives a general form for the characteristic function for any Lévy process. It assert that if we can describe the characteristic function of a process then we have sufficient information to define the process.

Theorem 3.1.9 (Lévy–Khintchine representation) *If ρ is infinitely divisible, then*

$$\chi(u) = e^{\psi(u)} \quad u \in \mathbb{R}^d, \tag{3.1.2}$$

where the characteristic exponent is given by

$$\psi(u) = \mathrm{i} < a, u > -\frac{1}{2} < u, bu > + \int_{\mathbb{R}^d} \left(e^{\mathrm{i}<u,x>} - 1 - \mathrm{i} < u, x > \mathbb{I}_{\{|x| \leq 1\}}(x)\right) \nu(dx),$$

where $a \in \mathbb{R}^d$, b is a positive-definite $d \times d$ matrix and ν is a measure on $\mathbb{R}^d - \{0\}$ satisfying

$$\nu(\{0\}) = 0, \text{ and satisfying integrability condition } \int_{\mathbb{R}^d} \left(\min\{|x|^2, 1\}\right) \nu(dx) < \infty.$$

Equation (3.1.2) *is uniquely defined by the characteristic triplet (a, b, ν). Conversely, for any value of a, b and ν satisfying above conditions which are required for a process to exhibit finite quadratic variation for a jump process to be semi-martingale, there exists an infinitely divisible distribution ρ having a representation in Equation 3.1.2.*

Stellenbosch University https://scholar.sun.ac.za

Proof. (see Sato, 2001, Theorem 1.3). □

If X_t is a Lévy process, then each X_t is infinitely divisible. To see this, take any $t > 0$. For any $n \in \mathbb{N}$, X_t can be expressed as

$$X_t = \sum_{i=1}^{n} \left(X_{\frac{it}{n}} - X_{\frac{(i-1)t}{n}} \right)$$

the sum of independent identical distributed (i.i.d) random variables.

Now using the stationary and independence of the increments, we can conclude that X_t is infinitely divisible ($\mathbb{P}_X = (\mathbb{P}_X)^n$). The characteristic function for a Lévy process X_t satisfies

$$\chi_X(u) = \mathbb{E}\left[e^{\mathrm{i}uX_t}\right] = \exp(t\psi(u)) = \exp\left\{ t \left(\mathrm{i}au - \frac{1}{2}ubu + \int_{\mathbb{R}^d} \left(e^{\mathrm{i}ux} - 1 - \mathrm{i}ux\mathbb{I}_{\{|x|\leq 1\}}(x)\right) \nu(dx) \right) \right\}.$$

Definition 3.1.10 (Lévy measure) *The Lévy measure ν on $\mathbb{R}^d$ must satisfy the following conditions:*

$$\nu(\{0\}) = 0 \;\; \textit{and} \;\; \int_{\mathbb{R}^d} \min\{1, |x|^2\} \, \nu(dx) < \infty.$$

The Lévy measure counts the expected number of jumps of a certain height in a particular time interval of length 1. In other words, let $A \in \mathcal{B}(\mathbb{R})$ be a Borel set in $\mathbb{R}$, the expected number of jumps of a particular size A in the time interval $[0, 1]$ is given by:

$$\nu(A) = \mathbb{E}\left[\#\{t \in [0, 1] : \Delta X_t \neq 0, \Delta X_t \in A\}\right].$$

Definition 3.1.11 *A random variable N_t is Poisson distributed if the probability for counting k jumps in the interval $[0, t]$ is equal to*

$$\mathbb{P}(N_t = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \;\; t > 0$$

with parameter $\lambda t = \mathbb{E}[N_t]$.

Definition 3.1.12 *Let $(\tau_i)_{i>1}$ be a sequence of independent exponential random variables with parameter λ, and $T_n = \sum_{i=1}^{n} \tau_i$. The process $(N_t, t \geq 0)$ defined by*

$$N_t = \sum_{n \geq 1} \mathbb{I}_{t \geq T_n}$$

is a Poisson process with intensity λ.

A Poisson process is a pure jump process such that the probability of more than one jump occurring in any sub-interval tends to zero. This process is too limited to develop realistic price models because its jumps are of constant size. It is sometimes required to have a process with random jump sizes. We can achieve this by giving some generalisation to a Poisson process by

Stellenbosch University https://scholar.sun.ac.za

letting $N_t = \sum_{i=1}^{N_t} 1$. Moreover, the process obtained by subtracting λt from a Poisson process is known as compensated Poisson process and it is denoted by

$$\tilde{N}_t = N_t - \lambda t.$$

The following definition applies:

Definition 3.1.13 (Compound Poisson process) *Let N_t be a Poisson process and Y_i independent and identically distributed $(i.i.d.)$ random variables that are independent of N_t. The process*

$$X_t = \sum_{i=1}^{N_t} Y_i, \quad t \in (0, \infty)$$

is called a compound Poisson process.

Although jumps in a compound process above happen at the same times as for the Poisson N_t, the $Y_i's$ are with non-unity jump size. This process is very important as they can approximate any Lévy process, i.e., Every infinitely divisible distribution is the limit of a sequence of compound Poisson distributions.

Define a set

$$E = \{(x, y) : x \in \mathbb{R}^+, y \in \mathbb{R}\}.$$

Consider $(x, t) \in E$, i.e. jump of size x at time t. A product measure μ on E is responsible for measuring the jump-size distribution. The measure μ_X for a compounded Poisson distribution is referred to as Poisson random measure.

Let $C \subseteq \mathbb{R}$ be a Borel set. Define a measure ν_X that counts the expected number of jumps in C with jump size

$$\nu_X(C) = \frac{\mathbb{E}[\mu_X(\cdot; [0, t], C)]}{t}.$$

The compensated compound Poisson process is given by

$$\tilde{X}_t = \int_0^t \int_{\mathbb{R}} x \tilde{\mu}_X(ds, dx),$$

where the compensated compound Poisson random measure is given by

$$\tilde{\mu}_X(\omega, [0, t], C) = \mu_X(\omega, [0, t], C) - t\nu_X(C).$$

Theorem 3.1.14 (Lévy–Itô decomposition) *Any Lévy process X_t can be represented in the following form,*

$$X_t = at + bW_t + \int_0^t \int_{\mathbb{R}} x \, \nu(ds, dx), \tag{3.1.3}$$

Stellenbosch University https://scholar.sun.ac.za

where a and $b \geq 0$ are real numbers and ν is a Lévy measure satisfying the usual conditions. $(W_t)_{t\geq 0}$ is a Brownian motion that is independent of ν. The third term is a compound Poisson process.

The Lévy–Itô decomposition says that a Lévy process has three parts:

(a) a deterministic part (drift) controlled by the drift parameter a.

(b) a Brownian motion part (diffusion) with parameter b.

(c) a pure jump with Lévy measure $\nu(u)$ that measures the intensity of jumps.

3.2 Path structure

The Lévy measure characterizes the path of the Lévy process in terms of activities and variation as follows:

Lemma 3.2.1 (Activity) *Let X_t be a Lévy process with Lévy triplet (a, b, ν) in one-dimension.*

*(i) If $\nu(\mathbb{R}) < \infty$ then almost all the paths of X_t have a finite number of jumps on every compact interval. We say the Lévy process has **finite activity**.*

*(ii) If $\nu(\mathbb{R}) = \infty$ then almost all the paths of X_t have an infinite number of jumps on every compact interval. We say the Lévy process has **infinite activity**.*

Lemma 3.2.2 (Variation) *Let X_t be a Lévy process with Lévy triplet (a, b, ν).*

*(i) If $b = 0$ and $\int_{|x|<1} |x|\ \nu(dx) < \infty$ then the paths of X_t has **finite variation** almost everywhere.*

*(ii) If $b \neq 0$ or $\int_{|x|<1} |x|\ \nu(dx) = \infty$ then the paths of X_t has **infinite variation** almost everywhere.*

The interested reader is referred to (Cont and Tankov, 2004, Chapter 3).

3.3 Generalised hyperbolic distribution

In this section we briefly discuss an example of Lévy process that we shall employ in later chapters for financial modeling.

Generalized hyperbolic (GH) distributions were first introduced by Barndorff–Nielsen(1977) with regard to modeling of grain size distribution of wind-blown sand. It was first applied

Stellenbosch University https://scholar.sun.ac.za

to financial modeling by Eberlein (2001) and Eberlein and Prause. The application of GH distribution became popular because of its ability to account for stylised features in financial return data for the underlying variable.

These distributions have five parameters. The one-dimensional Lebesgue density functions for generalised hyperbolic distributions are given by:

$$\rho_{\text{GH}}(x;\lambda,\alpha,\beta,\delta,\mu) = a(\lambda,\alpha,\beta,\delta)\left(\delta^2+(x-\mu)^2\right)^{\frac{(\lambda-\frac{1}{2})}{2}} K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)\exp\left(\beta(x-\mu)\right), \tag{3.3.1}$$

where α, β determine the shape of the distribution, δ is a scale factor, μ is responsible for the location, λ defines the tail fatness or classifies the distribution (see Section 6.4.1), and the constant factor a is defined by

$$a(\lambda,\alpha,\beta,\delta) = \frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^{\lambda}K_{\lambda}\left(\delta\sqrt{\alpha^2-\beta^2}\right)},$$

which is responsible for making the area under the curve equal to one and the function K_u is the modified Bessel function of the third kind with index u and it is given by

$$K_u(z) = \frac{1}{2}\int_0^{\infty} y^{u-1}\exp\left(-\frac{1}{2}z\left(y+\frac{1}{y}\right)\right)\,dy, \quad \text{for } z\in\mathbb{R}.$$

The domains of variation of the parameters are given in Table 3.1:

Table 3.1: GH parameter description

Θ	**Characteristics**	**Domain**
λ	Characterises the distribution	$\lambda \in \mathbb{R}$
α	Controls the behaviour of the tails	$\alpha \in \mathbb{R}^+$
β	Responsible for the skewness	$0 \leq \vert\beta\vert < \alpha$
δ	Scaling parameter (volatility)	$\delta \in \mathbb{R}^+$
μ	Responsible for the location	$\mu \in \mathbb{R}$

Source: Eberlein and Prause

These distributions are called generalised hyperbolic because their log-densities are hyperbolic whereas the log-density for a Gaussian distribution is a parabolic function.

These distributions have tails heavier than those for Gaussian distribution and have finite variance which can be approximated as follows:

$$\rho_{\text{GH}}(x;\lambda,\alpha,\beta,\delta,\mu) \sim |x|^{\lambda-1}\exp\left((\mp\alpha+\beta)x\right) \quad \text{as} \quad x\to\pm\infty. \tag{3.3.2}$$

Stellenbosch University https://scholar.sun.ac.za

The generalised hyperbolic distributions are proven to be closed under affine translation and parametrization, (see Eberlein, 2001). The former means if $X \sim GH(\lambda, \alpha, \beta, \delta, \mu)$, then $\tilde{X} = aX + b \sim GH(\lambda, \frac{\alpha}{|a|}, \frac{\beta}{a}, \delta|a|, a\mu + b)$ and the latter means that parameters can be parametrised as follows:

$$\xi = \frac{\beta}{\alpha}, \quad \zeta = \delta\sqrt{\alpha^2 - \beta^2}.$$

Generalized hyperbolic distribution is a class of distributions extensively used in finance modelling and is rich in structure. Two well-known subclasses are the normal inverse Gaussian (NIG) for $\lambda = -\frac{1}{2}$ and hyperbolic distribution (HYP) for $\lambda = 1$ by Barndorff-Nielsen (1998) and Eberlein and Keller (1995) respectively.

The normal inverse Gaussian (NIG) is the only special case for the GH which is closed under convolution, i.e. the sum of two independent normal inverse Gaussian distributed random variables is a normal inverse Gaussian distributed. It is because of this property that NIGs are widely used to price derivatives.

Following Eberlein (2001) closely, the generalised hyperbolic distribution can be represented as a mixture of a normal distribution with generalised inverse Gaussian (GIG). The probability density for GIG is given by

$$\rho_{\text{GIG}}(x; \lambda, \delta, \gamma) = \frac{\left(\frac{\delta}{\gamma}\right)^{\lambda}}{2K_{\lambda}(\sqrt{\gamma\delta})} X^{\lambda-1} \exp\left(-\frac{1}{2}(\delta^2 x + \gamma^2 x^{-1})\right). \tag{3.3.3}$$

GIGs are infinitely divisible. This is a necessary and sufficient condition for constructing a process. GIG distribution generates many subclass distributions. Two popular subclasses are inverse Gaussian distribution (IG) ($\lambda = -\frac{1}{2}$) and Gamma ($\delta = 0$).

As mentioned above, the density of GH can be expressed as a mixture of normal distribution with mean x and variance y ($\rho_{\text{N}}(x, y)$) and GIG, as follows:

$$\rho_{\text{GH}}(x; \lambda, \alpha, \beta, \delta, \mu) = \int_0^{\infty} \rho_{\text{N}}(x; \mu + \beta u, u)\rho_{\text{GIG}}(u; \lambda, \delta, \sqrt{\alpha^2 - \beta^2})\, du, \tag{3.3.4}$$

which is infinitely divisible since GIG is.

The GH Lévy measure has a closed form density given by

$$\nu_{\text{GH}}(x) = \begin{cases} \frac{e^{\beta x}}{|x|}\left(\int_0^{\infty} \frac{\exp(-\sqrt{2y+\alpha^2}|x|)}{\pi^2 y[J_{\lambda}^2(\delta\sqrt{2y})+K_{\lambda}^2(\delta\sqrt{2y})]}\, dy + \lambda e^{-\alpha|x|}\right), & \text{for } \lambda \geq 0, \\ \frac{e^{\beta x}}{|x|}\int_0^{\infty} \frac{\exp(-\sqrt{2y+\alpha^2}|x|)}{\pi^2 y[J_{-\lambda}^2(\delta\sqrt{2y})+K_{-\lambda}^2(\delta\sqrt{2y})]}\, dy, & \text{for } \lambda < 0, \end{cases} \tag{3.3.5}$$

where J_{λ} and K_{λ} are modified Bessel functions of the first and second kind respectively. The former is given by

$$J_{\lambda}(x) = \left(\frac{x}{2}\right)^{\lambda} \sum_{k=0}^{\infty} \frac{\left(\frac{x^2}{4}\right)^k}{k!\Gamma(\lambda + k + 1)}. \tag{3.3.6}$$

Stellenbosch University https://scholar.sun.ac.za

Using Theorem 3.1.2, the GH characteristic function is of the following form (see Eberlein and Prause):

$$\chi_{\text{GH}}(u) = e^{\psi(u)}, \quad \text{where} \quad \psi(u) = \mathrm{i}u\mathbb{E}[GH] + \int_{\mathbb{R}} \left(e^{\mathrm{i}ux} - 1 - \mathrm{i}ux\right) \nu_{\text{GH}}\, dx. \tag{3.3.7}$$

The analytical expression for this characteristic function is

$$\chi_{\text{GH}}(u) = e^{\mathrm{i}\mu u} \left(\frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + \mathrm{i}u)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda(\delta\sqrt{\alpha^2 - (\beta + \mathrm{i}u)^2}}{K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})}, \quad u \in \mathbb{R}, \tag{3.3.8}$$

which is a real-valued analytic function and can be extended to a holomorphic function along-side the strip

$$S := \{z : \beta - Im(z) < |\alpha|\}.$$

This is necessary for calculating the characteristic function in an extended manner, for instance for finding the moment generating function.

The moment generating function is obtained from the characteristic as follows:

$$M_{\text{GH}}(u) = \chi_{\text{GH}}(-\mathrm{i}u) = e^{\mu u} \left(\frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + u)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda(\delta\sqrt{\alpha^2 - (\beta + u)^2}}{K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})}, \quad |\beta + u| < \alpha.$$

This means that GH posses a moment of finite arbitrary order, which is necessary for derivative pricing. Hence we can find analytical expression for moments of any order. The formulas for the first two moments of a process X_t generated by GH are

$$\mathbb{E}[X_1] = \mu + \frac{\beta\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} \text{ and } Var[X_1] = \frac{\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} + \frac{\beta^2\delta^4}{\zeta^2}\left(\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \frac{K^2_{\lambda+1}(\zeta)}{K^2_\lambda}\right),$$

where $\zeta = \delta\sqrt{\alpha^2 - \beta^2}$.

The GH distributions are proven to be infinitely divisible. Hence, we can construct a GH process X, whose increments are of length 1. The Lévy process X_t generated by GH distribution is a pure jump process with Lévy triplet $(\mathbb{E}[X_t], 0, \nu_{\text{GH}}(dx))$ (see Eberlein, 2001). Furthermore, the GH process has paths of infinite activity.

3.4 Construction of a Lévy process

A common approach to constructing a Lévy process is to consider an arithmetic Brownian motion W_t and then change the flow of time from t to $\tau(t)$ for some stochastic process τ. This method is called time-changing standard Brownian motion or Brownian motion subordination. If τ is chosen to be a Lévy process then $W_{\tau(t)}$ is a Lévy process.

Stellenbosch University https://scholar.sun.ac.za

Definition 3.4.1 (Subordinator) *A stochastic process $\tau : \Omega \to \mathbb{R}$ is called a subordinator if it is an increasing process:*

$$\text{if } t_1 \leq t_2 \Longrightarrow \tau(t_1) \leq \tau(t_2) \ a.s.$$

Theorem 3.4.2 *A Lévy process X is a subordinator if and only if*

$$\nu(-\infty, 0) = 0, \int_{\mathbb{R}+} \min\{1, x\} \ \nu(dx) < \infty, b = 0 \ \text{ and } \ a = a - \int_{|x| \leq 1} x \ \nu(dx) \geq 0.$$

The GH process X_t is constructed using the subordinator for a drifted Brownian motion as follows:

$$X_t = \mu t + \beta\tau(t) + W_{\tau(t)}, \tag{3.4.1}$$

where $\tau(t)$ is GIG distributed with parameters λ, δ and $\sqrt{\alpha^2 - \beta^2}$.

The following theorem is vital for subordination.

Theorem 3.4.3 *(Cont and Tankov, 2004, Theorem 4.2) Let Y_t and τ be independent Lévy processes with Lévy characteristic exponents $\varphi(u)$ and $\upsilon(u)$ defined by triplets (a_Y, b_Y, ν_Y) and $(a_\tau, b_\tau, \nu_\tau)$. If τ is a subordinator then the process $X_t(\omega) = Y_{\tau_t(\omega)}(\omega)$ is a Lévy process and*

$$\mathbb{P}[X_t \in B] = \int_0^\infty \mu_Y^s(B)\mu_\tau^t(ds), \quad B \in \mathcal{B}(\mathbb{R}),$$

$$\phi_X(u) = \exp\left(\upsilon(-i\varphi(u))\right), \quad u \in \mathbb{R}.$$

The Lévy triplet (a_X, b_X, ν_X) for a process X_t is as follows:

$$a_X = a_\tau a_Y, \quad b_X = b_\tau b_Y + \int_0^\infty \nu_\tau(ds) \int_{\{|x| \leq 1\}} x\mu_Y^s(dx),$$

$$\nu_X = b_\tau \nu_Y(B) + \int_0^\infty \mu_Y^s(B)\nu_\tau(ds), \quad B \in \mathcal{B}(\mathbb{R} - \{0\}).$$

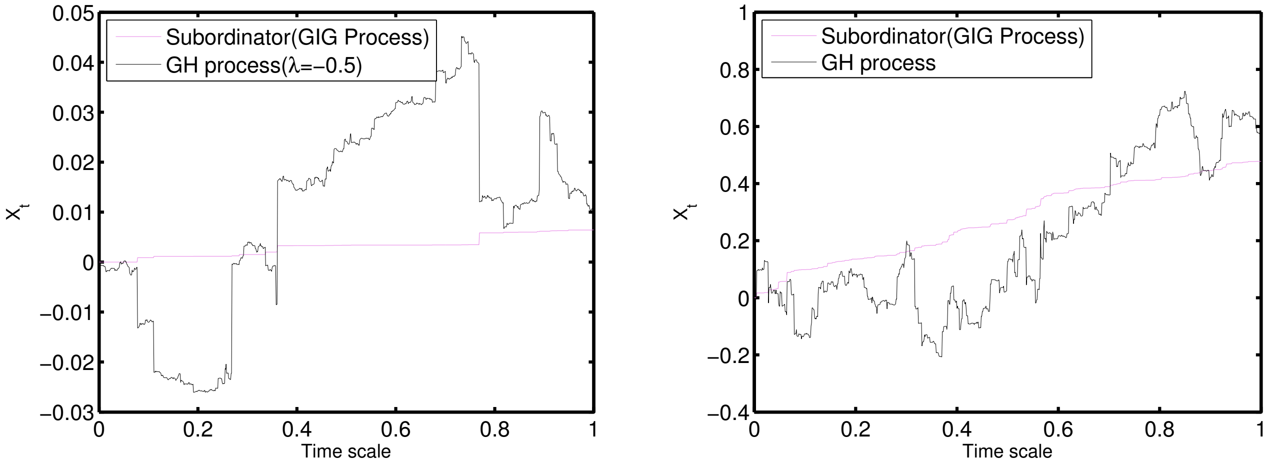
Algorithm simulation of GH process$(\lambda, \alpha, \beta, \delta, \mu)$

1. **for** $i \leftarrow 1$ **to** n
2. **do** $\Delta t_i = t_i - t_{i-1}$
3. $a \leftarrow (\delta\Delta t_i)^2$ and $b \leftarrow \alpha^2 - \beta^2$
4. simulate $I_i \sim GIG(\lambda, a, b)$[1]
5. simulate $W_i \sim N(0, 1)$
6. Compute $\Delta X_i \leftarrow \mu\Delta t_i + \beta I_i + \sqrt{I_i}W_i$
7. The GH discretised trajectory

$$X_{t_i} \leftarrow \sum_{j=1}^{i} \Delta X_j.$$

[1]GIG is a MATLAB coded program found in randraw.m.

Stellenbosch University https://scholar.sun.ac.za



(a) Simulation of NIG process, $\lambda = -0.5, \alpha = 15.1, \beta = -0.26, \delta = 0.09, \mu = 0$.

(b) Simulation of GH process, $\lambda = -0.02, \alpha = 15.1, \beta = -0.26, \delta = 0.09, \mu = 0$.

One can also construct a GH process by compound Poisson approximation (see Raible, 2000, Section 2.6.3).

3.5 Stochastic calculus

Definition 3.5.1 *An adapted stochastic process $X = (X_t)_{t \geq 0}$ is a semi-martingale if it admits the decomposition*

$$X = X_0 + M + A,$$

where X_0 is finite and $\mathcal{F}_t$- measurable, M is a local martingale that starts at zero, A is a càdlàg process with paths of finite variation and $A_0 = 0$. Furthermore, X is called a special semi-martingale if A is a predictable process.

Any special semi-martingale X has the following form

$$X_t = X_0 + W_t + X_t^c + \int_0^t \int_{\mathbb{R}} x(\mu - \nu)(ds, dx),$$

where X^c denotes the continuous part of X and the last term is the discontinuous part of X. μ is the random measure of the magnitude of jumps of X and ν is a stochastic compensator of μ.

Every Lévy process is a semi-martingale. Hence, it can be shown that a Lévy process X with Lévy triplet (a, b, ν) and satisfying $\mathbb{E}[X_1] < \infty$, has the following form:

$$X_t = at + \sqrt{b} W_t + \int_0^t \int_{\mathbb{R}} x(\mu - \nu)(ds, dx). \tag{3.5.1}$$

Stellenbosch University https://scholar.sun.ac.za

Here,

$$\int_0^t \int_{\mathbb{R}} x\mu(ds,dx) = \sum_{0\le s\le t} \Delta X_s \mathbb{I}_{\{|\Delta X_s|>1\}} \text{ and } \mathbb{E}\left[\int_0^t \int_{\mathbb{R}} x\mu(ds,dx)\right] = \int_0^t \int_{\mathbb{R}} x\nu(ds,dx) = t\int_{\mathbb{R}} x\,\nu(dx).$$

$\Delta X_s = X_s - X_{s^-}$ denotes the jump at time s. From the representation in Equation 3.5.1, if $a = 0$ (no drift term) then X_t is a martingale whereas if $b = 0$ then X_t is a purely discontinuous process.

Lemma 3.5.2 (Itô's Lemma) *Let $X = (X_t)_{t\geq 0}$ be a real-valued semi-martingale and f a C^2 function on $\mathbb{R}$. The function $f(X)$ is a semi-martingale and is given as follows:*

$$\begin{aligned} f(X_t) = f(X_0) &+ \int_0^t f'(X_{s_-})\, dX_s + \frac{1}{2}\int_0^t f''(X_s)\, ds \\ &+ \int_0^t \int_{\mathbb{R}} \left(f(X_{s_-}+x) - f(X_s) - xf'(X_{s_-})\right)\mu(ds,dx) \\ &= f(X_0) + \int_0^t \sqrt{b} f'(X_{s_-})\, dW_s + \int_0^t \int_{\mathbb{R}} \left(f(X_{s_-}+x) - f(X_s)\right)(\mu-\nu)(ds,dx) \\ &+ \int_0^t a f'(X_{s_-})\, ds + \frac{1}{2}\int_0^t b f''(X_s)\, ds \\ &+ \int_0^t \int_{\mathbb{R}} \left(f(X_{s_-}+x) - f(X_s) - xf'(X_{s_-})\mathbb{I}_{|x|\le 1}\right)\nu(ds,dx) \end{aligned} \tag{3.5.2}$$

Let $Y_t = f(X_t) = \exp(X_t)$. As a direct application of Itô's Lemma, the dynamics of Y_t are as follows:

$$\frac{dY_t}{Y_{t_-}} = \frac{1}{2}b\, dt + a\, dt + \sqrt{b}\, dW_t + \int_{\mathbb{R}} (e^x - 1)\,(\mu-\nu)(ds,dx) + \int_{\mathbb{R}} \left(e^x - 1 - x\mathbb{I}_{|x|\le 1}\right)\,\nu(dt,dx) \tag{3.5.3}$$

Theorem 3.5.3 *Suppose X is a Lévy process with characteristic exponent φ and $\mathbb{E}\left[e^{uX_t}\right] < \infty$. Then the process $M_t = (M_t)_{t\le 0}$ defined by*

$$M_t = \frac{e^{uX_t}}{e^{t\varphi(u)}}, \quad u \in \mathbb{R} \textit{ is a martingale}.$$

Proof. We must show that

$$\mathbb{E}\left[M_t|\mathcal{F}_s\right] = M_s \quad \forall 0 \le s < t.$$

Let X be a Lévy process with characteristic exponent φ. Then

$$\mathbb{E}\left[\frac{e^{uX_t}}{e^{t\varphi(u)}}\middle|\mathcal{F}_s\right] = e^{-t\varphi(u)}\mathbb{E}\left[e^{uX_t}\middle|\mathcal{F}_s\right] = e^{-t\varphi(u)}e^{uX_s}\mathbb{E}\left[e^{u(X_t-X_s)}\middle|\mathcal{F}_s\right] = e^{-t\varphi(u)}e^{uX_s}\mathbb{E}\left[e^{u(X_t-X_s)}\right]$$

$$= e^{-t\varphi(u)}e^{uX_s}e^{(t-s)\varphi(u)} = e^{uX_s}e^{-s\varphi(u)} = \frac{e^{uX_s}}{e^{s\varphi(u)}} = M_s.$$

$\square$

Stellenbosch University https://scholar.sun.ac.za

Theorem 3.5.4 (Girsanov transformation) *Let X_t be a Lévy process with the representation in Equation 3.5.1. Assume $\mathbb{P}$ and $\mathbb{Q}$ are equivalent probability measures, there is a deterministic process δ and a predictable process H such that*

$$\int_0^t \int_{\mathbb{R}} |x(H(s,x)-1)|\, \nu(dx)ds < \infty, \quad \int_0^t (b\cdot\delta_s^2)\, ds < \infty.$$

The density process $\gamma = (\gamma_t)_{t\geq 0}$ is defined as

$$\gamma_t = \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|\mathcal{F}_t\right] = \exp\left[\int_0^t \delta_s\sqrt{b}\, dW_s - \frac{1}{2}\int_0^2 \delta_s^2 b\, ds + \int_0^t\int_{\mathbb{R}}(H(s,x)-1)(\mu-\nu)(ds,dx)\right. \tag{3.5.4}$$
$$\left.-\int_0^t\int_{\mathbb{R}}(H(s,x)-1-\log(H(s,x)))\mu(ds,dx).\right]$$

If $\mathbb{E}[\gamma]=1$, then γ defines a probability measure $\mathbb{Q}$ such that $\mathbb{P}\sim\mathbb{Q}$.

Furthermore, the $\mathbb{Q}$-Brownian motion is given by

$$\hat{W}_t = W_t - \int_0^t \delta_s\, ds,$$

Hence under $\mathbb{Q}$, X_t has a representation

$$X_t = \hat{a}t + \sqrt{b}\hat{W}_t + \int_0^t\int_{\mathbb{R}} x(\mu-\hat{\nu})(ds,dx), \tag{3.5.5}$$

where

$$\hat{\nu}(ds,dx) = H(s,x)\nu(ds,dx) \;\; \textit{is the } \mathbb{Q} - \textit{compensator of } \mu$$

and

$$\hat{a}t = at + \int_0^t b\delta_s\, ds + \int_0^t\int_{\mathbb{R}} x(H(s,x)-1)\, \nu(ds,dx).$$

From the above theorem, we can deduce the Girsanov Theorem of Brownian motion whereby we set $\nu(dx)=0$ and $\delta_s = S(s,T)$ as follows:

Theorem 3.5.5 (Girsanov Theorem for Brownian motion) *Let W_t be a $\mathbb{P}$-Brownian motion. Suppose λ_t is a d-dimensional predictable process satisfying the Novikov condition*

$$\mathbb{E}_{\mathbb{P}}\left[\exp\left(\frac{1}{2}\int_0^T ||\lambda_s||^2\, ds\right)\right] < \infty,$$

then there is a martingale $\mathbb{Q}$ such that

(a) *$\mathbb{Q}$ is equivalent to $\mathbb{P}$*

(b) $\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\frac{1}{2}\int_0^T ||\lambda_s||^2\, ds - \int_0^T \lambda_s\, dW^{\mathbb{P}}\right), \quad \forall t\in[0,T]$

(c) $W_t^{\mathbb{Q}} = W_t - \int_0^t \lambda_s\, ds$ *is a $\mathbb{Q}$-Brownian motion*

Stellenbosch University https://scholar.sun.ac.za

3.6 Non-homogeneous Lévy processes

The Lévy processes discussed above are known as time-homogeneous Lévy processes. If we exclude the assumption of stationality of the increments we get non-homogeneous Lévy processes or time-inhomogeneous Lévy processes. Non-homogeneous Lévy processes are more general than homogeneous Lévy processes because of the flexibility of time-inhomogeneity they offer in pricing models Kluge (2005).

Definition 3.6.1 (Non-homogeneous Lévy process) *An adapted and càdlàg stochastic process $X = (X_t)_{t\geq 0}$ in $\mathbb{R}^d$ is called a non-homogeneous Lévy process if the following conditions hold:*

(a) X has increments independent of the past, i.e. $\forall\ 0 \leq s < t, \quad X_t - X_s \perp\!\!\!\perp \mathcal{F}_s$.

(b) The law of X_t, for all $t \in [0, T]$, is given via the characteristic function

$$\mathbb{E}\left[e^{\mathrm{i}<u,X_t>}\right] = \exp\int_0^t \left(\mathrm{i} < a_s, u > -\frac{1}{2} < u, b_s u > + \int_{\mathbb{R}^d} \left(e^{\mathrm{i}<u,x>} - 1 - \mathrm{i} < u, x > \mathbb{I}_{\{|x|\leq 1\}}(x)\right) \nu_t(dx \tag{3.6.1}$$

where $a_t \in \mathbb{R}^d$, b_t is a symmetric positive definite $d \times d$ matrix and ν_t is a Lévy measure on $\mathbb{R}^d$.

If

$$\int_0^{\mathbb{T}} \left(|a_s| + ||b_s|| + \int_{\mathbb{R}} \min\{|x|^2, 1\}\ \nu_s(dx)\right)\ ds < \infty,$$

then a non-homogeneous Lévy process is a semi-martingale often called a process with independent increments and absolutely continuous characteristics (PIIAC) (see Kluge, 2005, Lemma 1.4).

Example 3.6.2 (Non-homogeneous Lévy process) *Let $L = (L_t)_{t\geq 0}$ be a homogeneous Lévy process and $f(s)$ be a deterministic function. The process*

$$X_t = \int_0^t f(s)\ dL_s \tag{3.6.2}$$

is a non-homogeneous Lévy process (see Cont and Tankov, 2004, Example 14.4).

Processes such as in Equation 3.6.2 above are very important in financial modelling as they try to explain the time t portfolio value containing $f(s)$ risky assets and whose price follows an exponential Lévy process L_t. Thus, we are interested in describing this process for derivatives pricing purposes. Our first insight in describing the process is to find its explicit characteristic function. Using the same technique as in the above proof,

$$\chi_{X_t}(u) = \mathbb{E}\left[e^{iuX_t}\right] = \mathbb{E}\left[\exp\left(iu\int_0^t f(s)\ dL_s\right)\right] = \mathbb{E}\left[\exp\left(iu\sum_{k=0}^{n-1} f(t_k)[L(t_{k+1}) - L(t_k)]\right)\right],$$

Stellenbosch University https://scholar.sun.ac.za

by independent increments of L_t:

$$= \prod_{k=0}^{n-1} \mathbb{E}\left[\exp\left(iuf(t_k)[L(t_{k+1}) - L(t_k)]\right)\right]$$

and by stationality of increments of L_t:

$$= \prod_{k=0}^{n-1} \mathbb{E}\left[\exp\left(i(uf(t_k))L(\Delta t_k)\right)\right] = \prod_{k=0}^{n-1} \mathbb{E}\left[\exp\left((\psi(uf(t_k))\,\Delta t_k\right)\right] = \exp\left(\sum_{k=0}^{n-1} \psi(uf(t_k)\Delta t_k\right)$$
$$= \exp\left(\int_0^t \psi(uf(s))\,ds\right).$$

Lemma 3.6.3 *The characteristic function of a non-homogeneous Lévy in Equation 3.6.2 is given by:*

$$\chi_{X_t}(u) = \exp\left(\int_0^t \psi_s(uf(s))\,ds\right), \tag{3.6.3}$$

where $\psi(u) = \log \mathbb{E}[e^{iuL_1}]$.

The second main step is to simulate the process X_t that involves stochastic integrals. Let t_i be a regular partition of the interval $[0,t]$. Then the quadratic covariance of two semi-martingales f and L_t is given by

$$[f, L]_t = \lim \sum_{i=1}^{N} \left(f(t_{i+1}) - f(t_i)\right)\left(L_{i+1} - L_i\right)$$
$$= \lim \sum_{i=1}^{N} \left(f(t_{i+1})L_{t_{i+1}} - f(t_i)L_{t_i}\right) - \lim \sum_{i=1}^{N} f(t_i)\left(L_{i+1} - L_i\right) - \lim \sum_{i=1}^{N} L_{t_i}\left(f(t_{i+1}) - f(t_i)\right)$$
$$= f(t)L_t - f(0)L_0 - \int_0^t f(s_-)\,dL_s - \int_0^t L_{s_-}\,df(s).$$

Recall that paths of a Lévy process L_{s^-} and L_s are almost surely the same. Now since f is assumed to be continuous, $[f, L]_t = 0$ from the definition. Therefore, we have

$$\int_0^t f(s)\,dL_s = f(t)L_t - f(0)L_0 - \int_0^t L_s\,df(s)$$
$$= f(t)L_t - f(0)L_0 - \int_0^t L_s f'(s)\,ds. \tag{3.6.4}$$

The integral in Equation 3.6.4 can be performed fairly easily using Riemann sums. An alternative approach to simulating a process X_t is to realise that X_t is a Lévy process with different Lévy triplets which can by simulated in the usual way of simulating a Lévy process, i.e. approximating it with a compound Poisson process or Lévy–Itô decomposition.

Stellenbosch University https://scholar.sun.ac.za

Lemma 3.6.4 *A non-homogeneous Lévy process is an additive process in law.*

Proof. Suppose X is a non-homogeneous Lévy process. Since X has independent increments, it follows that for $0 \leq s \leq t \leq \mathbb{T}$,

$$\mathbb{E}\left[e^{\mathrm{i}<u,X_t>}\right] = \mathbb{E}\left[e^{\mathrm{i}<u,X_t-L_s>}\right]\mathbb{E}\left[e^{\mathrm{i}<u,X_s>}\right].$$

It is easy to see that $\mathbb{E}\left[e^{\mathrm{i}<u,X_t-X_s>}\right] = \exp\left(\int_s^t \psi_s(\mathrm{i}u)\, ds\right)$.
Observe that

$$\lim_{s\to t^-} \mathbb{E}\left[e^{\mathrm{i}<u,X_t-X_s>}\right] = 1, \quad \text{for fixed } u \in \mathbb{R}^d.$$

This means that

$$\lim_{s\to t} \mathbb{P}\left(|X_t - X_s| > \epsilon\right) = 0, \quad \text{for all } \epsilon > 0.$$

Hence, X_t is stochastically continuous and the characteristic function for X_0 is identically 1 and $X_0 = 0$. □

3.7 Exponential Lévy process

Throughout this section, by a Lévy process we mean a process as defined in Equation 3.6.2.

Definition 3.7.1 (Exponential Lévy process) *Let X be a Lévy process with Lévy characteristics $(a_t, b_t, \nu_t(d(x)))$ and satisfying*

$$\int_{|y|\geq 1} e^y \nu_t(dy) < \infty.$$

The process $Y_t = \exp(X_t)$ is an exponential Lévy process.

From the above definition, the Lévy–Itô decomposition of Y_t is $Y_t = M_t + A_t$, where

$$M_t = 1 + \int_0^t Y_{s^-} b_s\, dW_s + \int_0^t \int_{\mathbb{R}} Y_{s^-}(e^z - 1)\, \mu(ds, dz)$$

and

$$A_t = \int_0^t Y_{s^-}\left[a_t + \frac{b_t}{2} + \int_{\mathbb{R}} (e^z - 1 - z\mathbb{I}_{|z|\geq 1})\, \nu_t(dz)\right] ds.$$

The stochastic differential for an exponential Lévy process as in Definition 3.7.1 is

$$dY_t = Y_{s^-}\, dL_s, \tag{3.7.1}$$

where L is a Lévy process.

Stellenbosch University https://scholar.sun.ac.za

Theorem 3.7.2 (Doléans-Dade exponential) *Let X be a Lévy process with Lévy triplet $(a_t, b_t, \nu_t(dx))$. There exists a unique càdlàg process Y_t such that $dY_t = Y_{s^-}\, dL_s$ and $Y_0 = 1$. The process Y_t is called the Doléans-Dade exponential and has the following representation:*

$$Y_t = \varepsilon(X_t) = e^{X_t - \frac{b_t}{2}} \prod_{0<s\leq t} (1 + \Delta X_s) e^{-\Delta X_s}.$$

Furthermore, if $\int_{[-1,1]} |x|\ \nu_t(dx) < \infty$, then

$$Y_t = \varepsilon(X_t) = e^{b_t W_t + a_0 t - \frac{b_t}{2}} \prod_{0<s\leq t} (1 + \Delta X_s) \ \textit{where} \ \ a_0 = a_t - \int_{[-1,1]} |x|\ \nu_t(dx). \tag{3.7.2}$$

Proof. See (Cont and Tankov, 2004, Proposition 8.21). □

Stellenbosch University https://scholar.sun.ac.za

Chapter 4

Presentation of Lévy HJM models

In the last chapter we introduced Lévy processes and some basic results. Recall that Lévy processes are stochastic processes which offer excellent modeling features while also preserving analytical tractability. Lévy processes are processes with independent and stationary increments which turned out to be too restrictive for financial modeling. With stationary increments, it is proven that Lévy processes result in rigid scaling properties for the marginal distribution of the asset return which is not empirically observed in the time series, according to (Cont and Tankov, 2004, Chapter 14). Moreover, the shape of the caps/floor volatility surface is very complicated along the maturity axis to be reproduced by a model driven by a homogeneous Lévy process. This is mainly because models driven by a homogeneous Lévy process do not account for time inhomogeneity and are not structure preserving under change of measure. Consequently, a need arose to generalise the homogeneous Lévy model by considering non-homogeneous Lévy processes which are invariant under the change of measure and therefore simplify the computational complexity. These are processes with independent but not stationary increments. They are the generalised version of Lévy processes which provide more flexibility in the models. In other words, the use of non-homogeneous Lévy processes allows us to describe the dynamics of interest rates much better. As a result, they are crucial for the accurate calibration of interest rate models across different strikes and maturities.

Non-homogeneous Lévy processes, are sometimes referred to as time-inhomogeneous Lévy processes, are additive processes in law (see Lemma 3.6.4). Chapter 14 of the book of Cont and Tankov (2004) examine these processes in depth.

This chapter presents a general introduction the HJM framework as well as an extension in which a Brownian motion is replaced by a general non-homogeneous Lévy process. This framework is often referred to as an extended HJM framework or a Lévy HJM framework. In other words, we consider a HJM framework where the driver is a stochastic integral of a deterministic function with respect to a non-homogeneous Lévy process, which again form

Stellenbosch University https://scholar.sun.ac.za

a Lévy process (non-homogeneous). The main reason of applying a non-homogeneous Lévy process in a stochastic integral is that a driving process remains a non-homogeneous process after the change of measure and the model is analytically tractable, allowing for derivation of closed-form formulae.

To this end, we discuss the change of numéraire which form a strong basis for this study.

This study is based on the works of Björk *et al.* (1997), Eberlein and Raible (1999), Raible (2000), Eberlein *et al.* (2005) and a generalised study by Kluge (2005).

4.1 HJM methodology

The HJM approach is to consider the entire forward rate curve instead of the short rate as the fundamental quantity in the modelling of interest rates. Since we can deduce zero-coupon bond prices directly from the forward rate curve by Equation 2.2.2, we need to ensure the model does not give any opportunities for arbitrages. The HJM framework was a major breakthrough in modelling term structure, as it provides a necessary and sufficient condition for no-arbitrage in a model where the source of uncertainty is a Brownian motion. As a generalisation of the classical HJM model, Eberlein and Raible (1999) carried out an extension to a model where the sources of uncertainties are processes with jumps. Eberlein *et al.* (2005) and Filipović and Tappe (2008) gave conditions for no-arbitrage in HJM models driven by Lévy processes.

Since our main objective is to study a generalisation of the HJM framework, we begin with a short review of the underlying theory of the HJM framework.

Definition 4.1.1 (Heath *et al.* (1992)) *For any fixed maturity $T \leq \mathbb{T}$, the dynamics of the instantaneous forward rate $f(t,T)$ are given by an Itô process in differential form*

$$df(t,T) = \alpha(t,T)\, dt + \sigma(t,T)\, dW_t, \tag{4.1.1}$$

where W_t is assumed to be a d-dimensional $\mathbb{P}$-Brownian motion. Furthermore, it is assumed that

(a) the processes $\alpha(t,T)$ and $\sigma(t,T)$ are predictable processes which can both depend on the history of the Brownian motion and forward rates themselves up to time t, and are integrable, i.e

$$\int_0^T |\alpha(s,T)|\, ds + \int_0^T |\sigma(s,T)|^2\, ds < \infty \quad \mathbb{P}-a.s.$$

(b) the drift term α has finite integral $\int_0^T \int_0^u |\alpha(t,u)|\, dt\, du$

(c) the volatility σ has finite expectation, $\mathbb{E}\left[\int_0^T |\int_0^u \alpha(t,u)\, dW_t|\, du\right]$

Stellenbosch University https://scholar.sun.ac.za

(d) *the initial forward rate curve* $f(0, T)$ *is deterministic and* $\int_0^T |f(0, u)|\, du < \infty$.

The first and last assumptions in the above theorem ensure that the forward rates $f(t, T)$ are well-behaved and defined by their stochastic differential equation while the second and third assumptions are necessary for Fubini's theorem, which asserts that the stochastic differential of the integral of $f(t, T)$ with respect to T is the same as the integral of the stochastic differentials of $f(t, T)$.

Consequently, the dynamics of the forward rates in Equation 4.1.1 can be expressed as

$$f(t, T) = f(0, T) + \int_0^t \alpha(s, T)\, ds + \int_0^t \sigma(s, T)\, dW_s \ \ \forall t \in [0, T], \tag{4.1.2}$$

where $f(0, T)$ are the initial forward rates (-known at the present time) taken from the market instantaneous forward curve $T \mapsto f(0, T)$.

To obtain the bond dynamics, substitute Equation 4.1.2 into the relation in Equation 2.2.2.

$$\begin{aligned} P(t, T) &= \exp\left(-\int_t^T \left(f(0, u) + \int_0^t \alpha(s, u)\, ds + \int_0^t \sigma(s, u)\, dW_s\right) du\right) \\ &= \exp\left(-\int_t^T f(0, u)\, du - \int_t^T \int_0^t \alpha(s, u)\, ds\, du - \int_t^T \int_0^t \sigma(s, u)\, dW_s\, du\right) \\ &= \exp\left(-\int_t^T f(0, u)\, du\right) \times \exp\left(-\int_t^T \int_0^t \alpha(s, u)\, ds\, du - \int_t^T \int_0^t \sigma(s, u)\, dW_s\, du\right) \\ &= \frac{P(0, T)}{P(0, t)} \exp\left(-\int_t^T \int_0^t \alpha(s, u)\, ds\, du - \int_t^T \int_0^t \sigma(s, u)\, dW_s\, du\right). \end{aligned} \tag{4.1.3}$$

Based on our assumptions, the function $\alpha(\cdot, T)$ and $\sigma(\cdot, T)$ are continuous on the region $R = [t, T] \times [0, t]$ see (Björk, 2009, Chapter 22). This enables us to apply stochastic Fubini theorem for iterated integrals. Recall from basic calculus that

$$\frac{d}{dt}\left(-\int_t^T f(t, u)\, du\right) = f(t, t) - \int_t^T \frac{\partial}{\partial t}\, df(t, u)\, du.$$

$$d\left(-\int_t^T f(t, u)\, du\right) = r_t\, dt - \int_t^T (\alpha(t, u)\, dt + \sigma(t, u)\, dW_t)\, du = r_t dt + \int_t^T -\alpha(t, u)\, dudt + \int_t^T -\sigma(t, u)\, dudW_t$$

Let $\alpha^*(t, T) = \int_t^T \alpha(t, u)\, du$ and $\sigma^*(t, T) = \int_t^T \sigma(t, u)\, du$. Hence we have

$$d\left(-\int_t^T f(t, u)\, du\right) = r_t\, dt - \alpha^*(t, T)\, dt - \sigma^*(t, T)\, dW_t.$$

Stellenbosch University https://scholar.sun.ac.za

Apply the Itô formula to the relation in Equation 2.2.2.

$$\begin{aligned} dP(t,T) &= d\left(\exp\left(-\int_t^T f(t,u)\,du\right)\right) = \exp\left(-\int_t^T f(t,u)\,du\right) d\left(-\int_t^T f(t,u)\,du\right) \\ &\quad + \frac{1}{2}\left(\exp\left(-\int_t^T f(t,u)\,du\right)\left[d\left(-\int_t^T f(t,u)\,du\right)\right]^2\right) \\ &= P(t,T)\left[r_t\,dt - \alpha^*(t,T)\,dt - \sigma^*(t,T)\,dW_t\right] + \frac{1}{2}P(t,T)|\sigma^*(t,T)|^2\,dt \\ &= P(t,T)\left[r_t - \alpha^*(t,T) + \frac{1}{2}|S(t,T)|^2\right]\,dt - \sigma^*(t,T)P(t,T)\,dW_t. \end{aligned} \tag{4.1.4}$$

Therefore the bond dynamics are given by

$$dP(t,T) = P(t,T)\left[A(t,T)\,dt + S(t,T)\,dW_t\right], \tag{4.1.5}$$

where $A(t,T) = r_t - \alpha^*(t,T) + \frac{1}{2}|\sigma^*(t,T)|^2$, $S(t,T) = -\sigma^*(t,T)$ are the drift and volatility for the bond prices respectively. Furthermore, if the forward volatility function $\sigma(t,T)$ is a deterministic function, this case is known as Gaussian HJM.

The dynamics in Equations 4.1.1, 4.1.2 and 4.1.5 do not guarantee the absences of arbitrage. Heath *et al.* (1992) proved that in order to make sure that there is a martingale measure, and therefore no arbitrage, the drift $\alpha(t,T)$ function must be chosen in a special form. That is to say that the drift term $\alpha(t,T)$ cannot be chosen arbitrarily.

Let $Z(t,T) = \frac{P(t,T)}{B_t}$ be the discounted bond prices. We use Itô's formula to find the dynamics of discounted bond processes:

$$\begin{aligned} dZ(t,T) &= \frac{1}{B_t}\,dP(t,T) - \frac{P(t,T)}{B_t^2}\,dB_t \\ &= \frac{1}{B_t}\left(P(t,T)\left[A(t,T)\,dt + S(t,T)\,dW_t\right]\right) - \frac{1}{B_t}P(t,T)r_t\,dt \\ &= Z(t,T)\left(\left[A(t,T) - r_t\right]\,dt + S(t,T)\,dW_t\right) \\ &= Z(t,T)\left[\left(-\alpha^*(t,T) + \frac{1}{2}|\sigma^*(t,T)|^2\right)\,dt + S(t,T)\,dW_t\right]. \end{aligned} \tag{4.1.6}$$

Now under a risk-neutral measure $\mathbb{Q}$ we need a new Brownian motion, which can be obtained via Girsanov's theorem for Brownian motion in Theorem 3.5.5. If we define a d-dimensional process λ_t such that the Novikov condition is fulfilled:

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T ||\lambda_s||^2\,ds\right)\right] < \infty$$

and define $\mathbb{Q}$ so that the Radon–Nikodym derivative is given by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\int_0^t \frac{1}{2}\lambda_s^2\,ds - \int_0^t \lambda_s\,dW^{\mathbb{Q}}\right), \quad \forall t \in [0,T],$$

Stellenbosch University https://scholar.sun.ac.za

then

$$W_t^{\mathbb{Q}} = W_t - \int_0^t \lambda_s \, ds,$$

or in differential form

$$dW_t^{\mathbb{Q}} = dW_t - \lambda_t dt$$

is a $\mathbb{Q}$-Brownian motion.

Hence, under $\mathbb{Q}$, the discounted bond dynamics become

$$dZ(t,T) = Z(t,T)\left(\left[-\alpha^*(t,T) + \frac{1}{2}|\sigma^*(t,T)|^2 + S(t,T)\lambda_t\right] \, dt + S(t,T) \, dW_t^{\mathbb{Q}}\right).$$

We choose λ_t in such a way that the drift term vanishes, i.e.,

$$-\alpha^*(t,T) + \frac{1}{2}|\sigma^*(t,T)|^2 + S(t,T)\lambda_t = 0.$$

The process λ_t is often called market price of risk. It is the compensation paid per unit of volatility, or (risk taken).

Differentiation of the last equation with respect to T yields

$$\alpha(t,T) = \sigma(t,T)\sigma^*(t,T) - \sigma(t,T)\lambda_t = \sigma(t,T)\left(\sigma^*(t,T) - \lambda_t\right).$$

Hence, under the risk-neutral measure $\mathbb{Q}$ we have the following relation of the forward drift term and the volatility:

$$\alpha(t,T) = \sigma(t,T)\left(\sigma^*(t,T) - \lambda\right), \quad \forall t \in [0,T]. \tag{4.1.7}$$

Since under $\mathbb{Q}$ the drift term of Equation 4.1.5 becomes $A(t,T) + \lambda_t S(t,T)$, which is equal to

$$A(t,T) + S(t,T)\lambda_t = r_t \overbrace{-\alpha^*(t,T) + \frac{1}{2}|\sigma^*(t,T)|^2 + S(t,T)\lambda_t}^{=0} = r_t,$$

hence the zero-coupon bond dynamics under an equivalent martingale measure $\mathbb{Q}$ are given by

$$dP(t,T) = P(t,T)\left[r_t \, dt + S(t,T) \, dW_t^{\mathbb{Q}}\right], \tag{4.1.8}$$

where $r_t = f(t,t)$ is the instantaneous short rate at time t,

$$r_t = f(t,t) = \int_0^t \alpha(s,t) \, ds + \int_0^t \sigma(s,t) \, dW_s.$$

Hence, we have the HJM drift condition.

Stellenbosch University https://scholar.sun.ac.za

Theorem 4.1.2 (HJM drift condition) *There exists an equivalent martingale measure if and only if the drift term of the forward rate is of the form*

$$\alpha(t,T) = \sigma(t,T)\left(\int_t^T \sigma(t,u)\,du - \lambda_t\right) \quad \forall T \in [0,\mathbb{T}].$$

The HJM drift condition restricts the drift of the forward rates to prevent arbitrage opportunities in the market. Furthermore, modelling under a risk-neutral measure $\mathbb{Q}$ implies the particular choice of $\lambda_t = 0$ in Theorem 4.1.2.

In any term structure model satisfying the HJM drift condition in Theorem 4.1.2 the forward rate and the zero-coupon bond dynamics must evolve according to the following:

$$df(t,T) = \alpha(t,T)\,dt + \sigma(t,T)\,dW_t^{\mathbb{Q}} \text{ and } dP(t,T) = r_t\,dt + S(t,T)\,dW_t^{\mathbb{Q}}.$$

Hence the model is fully specified by the volatility structure $\{\sigma(t,T)\}_{T\geq t}$ and by the initial forward rates $f(0,T)$ observed from the market.

The short rate process can be rewritten as follows:

$$r_t = f(0,t) + \int_0^t \sigma(u,t) \int_u^t \sigma(u,s)ds\,du + \int_0^t \sigma(s,t)dW_s^{\mathbb{Q}} \tag{4.1.9}$$

Notice that under the risk-neutral measure $\mathbb{Q}$ the short rate process and forward rate dynamics both depend on the volatility function $\sigma(\cdot,T)$. Thus, in a short rate model we only specify a single volatility σ to solve for the forward rate volatility $\sigma(t,T)$. In HJM models, we begin by specifying the entire forward volatility surface.

In the the following, we summarise the condition for completeness of HJM framework.

Condition 4.1.3 (Market completeness) *(a) There is an adapted process λ_t such that*

$$\alpha(t,T) = \sigma(t,T)\left(\int_t^T \sigma(t,u)\,du - \lambda_t\right), \quad \textit{for all } t \leq T.$$

(b) The process $A(t,T) = \int_t^T \sigma(t,u)\,du$ is non-negative for every $t < T$.

(c)

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T ||\lambda_s||^2\,ds\right)\right] < \infty.$$

(d)

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T ||\int_t^T \sigma(t,u)\,du - \lambda_s||^2\,ds\right)\right] < \infty.$$

Stellenbosch University https://scholar.sun.ac.za

4.2 Markovianity in the HJM framework

Although the HJM framework provides an excellent framework for modeling term structure, it produces models in which short rates are path-dependent. Another drawback is that most HJM models are non-Markovian. For instance, the short rate implied from the forward rate in Equation 4.1.9 contains a variable t inside the stochastic integral as upper limit and also inside the integrand function, making the term structure evolution reliant on the path taken. Ultimately, zero-coupon bonds may be path-dependent, which may make it inefficient for computation as this means that Monte Carlo simulation techniques must be used to value options under these models. However, since forward rate dynamics are governed by the volatility structure, we can impose conditions on the volatility function which ensure that the short rate process implied from the HJM models are Markovian see (Carverhill, 1994; Ritchken and Sankarasubramanian, 1995).

Definition 4.2.1 (Markov property) *A stochastic process X satisfies the Markov property with respect to the risk-neutral measure $\mathbb{Q}$ if for all $s \leq t \leq \mathbb{T}$ and for any bounded and measurable function (Borel) $f : \mathbb{R} \mapsto \mathbb{R}$,*

$$\mathbb{E}_{\mathbb{Q}}\left[f(X_t)|\mathcal{F}_s\right] = \mathbb{E}_{\mathbb{Q}}\left[f(X_t)|X_s\right].$$

We often say that a process X_t is Markovian if it satisfies the Markov property above.

For a model driven by a Brownian motion or a Gaussian process, the short rate process is Markovian if and only if the volatility structure has a form of factorisation (see Carverhill (1994)), while a general form for volatility that guarantees the Markovian property of short rate process was stated and proven by Ritchken and Sankarasubramanian (1995). For a general Lévy process the same result was proved in the original paper of Eberlein and Raible (1999), where some conditions on the characteristic function associated with the Lévy process were imposed.

To summarise this, Theorem 4.2.2 will help us choose a volatility function $\sigma(t,T)$ that will ensure that the short rate inferred from forward rates is Markovian. The first version of this result was stated and proved in Ritchken and Sankarasubramanian (1995).

Theorem 4.2.2 *If the forward rate volatility structure is of the form*

$$\sigma(t,T) = \sigma_1(t)\cdot\sigma_2(T) \quad \textit{for all} \quad t \leq T \leq \mathbb{T},$$

where σ_1 and σ_2 are strictly positive, deterministic functions of time and continuously differentiable functions, then the short rate process r is a Markov process.

Proof. (See Ritchken and Sankarasubramanian, 1995, p.4). □

Stellenbosch University https://scholar.sun.ac.za

The converse of the above theorem is also true (see Eberlein and Raible, 1999, p.41) for a more generalised proof.

In this study, we consider the Vasiček volatility structure form given by

$$\sigma(t,T) = \hat{\sigma} e^{-a(T-t)} \quad \forall a, \hat{\sigma} \in \mathbb{R} \text{ and } \forall t \leq T \leq \mathbb{T}, \tag{4.2.1}$$

where a and $\hat{\sigma}$ represent the volatility and the mean-reversion parameter respectively of the Vasiček model.

To see that the conditions of Theorem 4.2.2 are satisfied, choose $\sigma_1 : [0,\mathbb{T}] \mapsto \mathbb{R}$ and $\sigma_2 : [0,\mathbb{T}] \mapsto (0,\infty)$ defined by

$$\sigma_1(t) = \hat{\sigma} e^{at} \text{ and } \sigma_2(T) = e^{-aT}.$$

The resulting forward rates dynamics are given by

$$f(t,T) = f(0,T) + \frac{\hat{\sigma}^2}{a} \int_0^t e^{-a(T-u)} \left(1 - e^{-a(T-u)}\right) du + \hat{\sigma} \int_0^t e^{-a(T-u)} \, dW_u^{\mathbb{Q}}$$

and the short rate process (4.1.9) becomes

$$\begin{aligned} r(t) &= f(0,t) + \sigma_2(t) \int_0^t \sigma_1^2(u) \int_u^t \sigma_2(s) \, ds \, du + \sigma_2(t) \int_0^t \sigma_1(s) \, dW_s^{\mathbb{Q}} \\ &= f(0,t) + \hat{\sigma}^2 e^{-at} \int_0^t e^{2au} \int_u^t e^{-as} \, ds \, du + \hat{\sigma} e^{-at} \int_0^t e^{as} \, dW_s^{\mathbb{Q}} \\ &= f(0,t) - \frac{\hat{\sigma}^2}{a} e^{-at} \int_0^t e^{2au} \left(e^{-at} - e^{-au}\right) du + \hat{\sigma} e^{-at} \int_0^t e^{as} \, dW_s^{\mathbb{Q}} \\ &= f(0,t) + \frac{\hat{\sigma}^2}{2a^2} \left(1 - e^{-at}\right)^2 + \hat{\sigma} e^{-at} \int_0^t e^{as} \, dW_s^{\mathbb{Q}}. \\ &= \mu(t) + \hat{\sigma} e^{-at} \int_0^t e^{as} \, dW_s^{\mathbb{Q}}. \end{aligned} \tag{4.2.2}$$

4.3 The driving process

Throughout this section, we assume that $L = (L_t)_{t \geq 0}$ is a one-dimensional non-homogeneous Lévy process adapted on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, with Lévy triplet (a_t, b_t, ν_t) (see also Cont and Tankov, 2004, Theorem 14.1). Since every Lévy process is a semi-martingale, its canonical representation is given by

$$L_t = \int_0^t a_s \, ds + \int_0^t \sqrt{b_s} \, dW_s + \int_0^t \int_{\mathbb{R}} x(\mu^L - \nu^L)(ds, dx), \tag{4.3.1}$$

Stellenbosch University https://scholar.sun.ac.za

where $(W_t)_{t\geq 0}$ is a d-dimensional standard Brownian motion, μ^L is the random measure of jumps of L and ν^L is the compensator of μ^L defined by

$$\nu^L(ds, dx) = \nu_t(dx)ds,$$

(see Section 3.1).

The HJM model driven by a one-dimensional stochastic process L_t for $t \leq \mathbb{T}$ with independent increments and absolutely continuous characteristics (PIIC) was first studied by Kluge (2005), who gave an empirical motivation for using these types of processes by showing that a model driven by non-homogeneous Lévy processes can reproduce the volatility surface.

Replacing a Brownian motion in Equation 4.1.1, the dynamics of the forward rate driven by L_t is given by

$$df(t,T) = \alpha_{\text{Lévy}}(t,T)dt - \sigma(t,T)dL_t \quad \forall(0 \leq t \leq T \leq \mathbb{T}), \tag{4.3.2}$$

where $\alpha_{\text{Lévy}}(t,T)$ and $\sigma(t,T)$ are deterministic functions whose paths are continuously differentiable with respect to T and the initial forward rates $f(0,T)$ are deterministic, bounded and measurable functions in T. Recall that $\alpha_{\text{BM}}(t,T)$ is a special case of $\alpha_{\text{Lévy}}(t,T)$. The minus sign in Equation 4.3.2, is put there for ease of calculation.

For every t the law of L_t is governed by the characteristic function

$$\mathbb{E}[e^{iuL_t}] = e^{\int_0^t \psi_s(\mathrm{i}u)\, ds} \text{ for every } t \in [0,\mathbb{T}], \tag{4.3.3}$$

where ψ represents the characteristics exponent for L_t as given by the Lévy–Khintchine Theorem 3.1.9 with Lévy triplet (a_t, b_t, ν_t).

The following assumptions are enforced to guarantee integrability.

Condition 4.3.1 *(Integrability)*

(I) There exist M*,* $\epsilon > 0$ *such that*

$$\int_0^{\mathbb{T}} \int_{\{|x|>1\}} e^{ux}\, \nu_s(dx) < \infty \quad \forall |u| \leq (1+\epsilon)M.$$

(II)

$$\int_0^{\mathbb{T}} \left(|a_s| + ||b_s|| + \int_{\mathbb{R}} \min\{|x|^2, 1\}\, \nu_s(dx) \right) ds < \infty.$$

The above conditions imply that L_t is a special semi-martingale, (see Kluge, 2005, Lemma 1.7) with canonical decomposition of the form Equation 4.3.1.

Since zero-coupon bonds are contained in the forward rates structure, the zero-coupon bond price process under Lévy settings is given as follows:

Stellenbosch University https://scholar.sun.ac.za

$$P(t,T) = \exp\left(-\int_t^T f(t,u)\,du\right),$$

which can also be expressed as

$$P(t,T) = P(0,T)B_t \exp\left(-\int_0^t A(s,T)\,ds + \int_0^t S(s,T)\,dW_s\right), \tag{4.3.4}$$

where B_t is the bank account in Equation 2.1.1 and $S(s,T) = \int_{s\wedge T}^T \sigma(t,u)\,du$, $A(s,T) = \int_{s\wedge T}^T \alpha(t,u)\,du$, where $s \wedge T = \min\{s,T\}$.

Theorem 4.3.2 *Suppose $f : \mathbb{R} \mapsto \mathbb{C}$ is a left continuous function with limits from the right such that $|\Re\{f(x)\}| \leq M$, $x \in dom(f)$, then*

$$\mathbb{E}\left[\exp\left(\int_0^t f(s)\,dL_s\right)\right] = \exp\left(\int_0^t \psi_s(f(s))\,ds\right).$$

Proof.

The characteristic function for any Lévy process is given by Lévy–Khintchine Theorem 3.1.9, which is also given in 6.4.1. Recall that a process $Y_t = \int_0^t f(s)\,dL_s$ is a process with independent increments, i.e. a non-homogeneous process see (Eberlein and Raible, 1999, Lemma 3.1). □

Replace a Brownian motion in Equation 4.3.4 by a Lévy process to get

$$P(t,T) = P(0,T)B_t \exp\left(-\int_0^t A(s,T)\,ds + \int_0^t S(s,T)\,dL_s\right). \tag{4.3.5}$$

Equation 4.3.5 has an equivalent form (see also Section A.1).

$$P(t,T) = P(0,T)B_t \frac{\exp\left(\int_0^t S(s,T)\,dL_s\right)}{\mathbb{E}\left[\exp\left(\int_0^t S(s,T)\,dL_s\right)\right]}. \tag{4.3.6}$$

The discounted bond prices are

$$Z(t,T) = \frac{P(t,T)}{B_t} = P(0,T)\exp\left(-\int_0^t A(s,T)\,ds + \int_0^t S(s,T)\,dL_s\right). \tag{4.3.7}$$

Define a process $Y(t,T) = -\int_0^t A(s,T)\,ds + \int_0^t S(s,T)\,dL_s$. The process $Y(t,T)$ is a semi-martingale and it admits the Lévy-Itô decomposition:

$$\begin{aligned} Y(t,T) = &-\int_0^t A(s,T)\,ds + \int_0^t a_s S(s,T)\,ds + \int_0^t \sqrt{b_s} S(s,T)\,dW_s \\ &+ \int_0^t \int_{\mathbb{R}} S(s,T)x(\mu^L - \nu^L)\,(dt,dx). \end{aligned} \tag{4.3.8}$$

Stellenbosch University https://scholar.sun.ac.za

Now $Z(t,T) = P(0,T)\exp(Y(t,T))$. Using Itô's Lemma 3.5.2 and Equation 3.5.3;

$$\begin{aligned}\frac{dZ(t,T)}{Z(t-,T)} &= \left(-A(s,T) + S(t,T)a_t + \frac{1}{2}b_t\,||S(t,T)||^2\right)dt + S(t,T)\sqrt{b_t}\,dW_t \\ &+ \int_{\mathbb{R}/\{0\}}\left(e^{S(t,T)x}-1\right)(\mu-\nu)\,(dt,dx) + \int_{\mathbb{R}/\{0\}}\left(e^{S(t,T)x}-1-S(t,T)x\right)\ \nu(dt,dx) \\ &= \left(-A(s,T) + S(t,T)a_t + \frac{1}{2}b_t\,||S(t,T)||^2 + \int_{\mathbb{R}/\{0\}}\left(e^{S(t,T)x}-1-S(t,T)x\right)\ \nu_t(dx)\right)\ dt \\ &+ S(t,T)\sqrt{b_t}\,dW_t + \int_{\mathbb{R}/\{0\}}\left(e^{S(t,T)x}-1\right)(\mu-\nu)\,(dt,dx).\end{aligned} \tag{4.3.9}$$

To avoid arbitrage opportunities, we ought to model discounted prices that form martingale. To make discounted prices above martingale, we must eliminate the drift term. That is to say

$$-A(s,T) + S(t,T)a_t + \frac{1}{2}b_t\,||S(t,T)||^2 + \int_{\mathbb{R}/\{0\}}\left(e^{S(t,T)x}-1-S(t,T)x\right)\ \nu_t(dx) = 0$$

$$A(t,T) = S(t,T)a_t + \frac{1}{2}b_t\,||S(t,T)||^2 + \int_{\mathbb{R}/\{0\}}\left(e^{S(t,T)x}-1-S(t,T)x\right)\ \nu_t(dx) = \psi_t(S(t,T)).$$

This leads us to the Lévy equivalent HJM drift condition.

Theorem 4.3.3 (Drift condition) *Under a risk-neutral measure $\mathbb{Q}$, the Lévy market model is arbitrage-free if we choose the drift term of the forward rates to be*

$$\alpha_{\text{Lévy}}(t,T) = \sigma(t,T)\psi_2\left(S(t,T)\right)$$

and

$$A(t,T) = \psi_t(S(t,T)),\ \forall T \in [0,T],$$

where ψ_t is the characteristic exponent associated with a Lévy process L_t defined by its Lévy triplet (a_t, b_t, ν_t), i.e., for all $|u| \le (1+\epsilon)M$,

$$\psi_t(u) = a_t u - \frac{1}{2}ub_t u + \int_{\mathbb{R}^d}\left(e^{ux}-1-ux\right)\nu_t(dx), \tag{4.3.10}$$

where $b_t = b_t + \int_{\mathbb{R}} x\mathbb{I}_{\{|x|>1\}}\ \nu_t(dx)$.

Notice that the truncation function $\mathbb{I}_{\{|x|\le 1\}}(x)$ does not appear in Equation 4.3.10, because of the conditions for integrability.

We have shown that discounted bond prices are martingale. In fact, Eberlein *et al.* (2005) showed that this risk-neutral martingale measure is unique.

The forward dynamics in Equation 4.3.2 merely depends on the volatility. We consider the deterministic volatility structures.

Stellenbosch University https://scholar.sun.ac.za

Condition 4.3.4 (Deterministic volatility) *We assume that the volatility structure σ is deterministic and bounded on $\Delta = \{(t,T) : 0 \le t, T \le \mathbb{T}\}$. That is to say,*

$$0 \le S(t,T) \le M,$$

where M satisfies the integrability conditions in 4.3.1 above. Furthermore, it is assumed that $S(t,T)$ is twice continuously differentiable function in both variables.

Let $S(t,T)$ be f in Theorem 4.3.2, i.e

$$S(t,T) = f(t), \quad \text{for fixed } T \in [0, \mathbb{T}].$$

Theorem 4.3.2 says that

$$\begin{aligned}\mathbb{E}\left[\exp\left(\int_0^t S(s,T)\, dL_s\right)\right] &= \exp\left(\int_0^t \left(S(s,T)a_s + \frac{1}{2}b_s||S(s,T)||^2 + \int_{\mathbb{R}} \left(e^{S(s,T)x} - 1 - S(s,T)x\right)\nu_s(dx)\right)\, ds\right) \\ &= \exp\left(\int_0^t \psi_s(S(t,T))\, ds\right) \\ &= \exp\left(\int_0^t A(s,T)\, ds\right). \end{aligned} \tag{4.3.11}$$

This means that the expectation in the denominator of Equation 4.3.6 is well defined. By substitution one gets

$$P(t,T) = P(0,T)B_t \exp\left(-\int_0^t \psi_t\left(S(s,T)\right)\, ds + \int_0^t S(s,T)\, dL_s\right). \tag{4.3.12}$$

The discounted bond prices in Equation 4.3.7 can be rewritten as

$$Z(t,T) = P(0,T)\frac{\exp\left(\int_0^t S(s,T)\, dL_s\right)}{\mathbb{E}\left[\exp\left(\int_0^t S(s,T)\, dL_s\right)\right]}. \tag{4.3.13}$$

See Apendix A for a detailed illustration how we arrive at Equation 4.3.13.

Since the volatility structure is assumed to be deterministic, we have the following from Equation 3.6.2:

Corollary 4.3.5 *Define*

$$Y_t := \int_0^t S(s,T)\, dL_s.$$

Then Y_t is a non-homogeneous process and its Lévy–Itô decomposition (following from Equation 4.3.1) is given by

$$Y_t = \int_0^t a_s S(s,T)\, ds + \int_0^t \sqrt{b_s} S(s,T)\, dW_s + \int_0^t \int_{\mathbb{R}} S(s,T)x\tilde{\mu}_s(ds,dx),$$

Stellenbosch University https://scholar.sun.ac.za

where a_s and $b_s \geq 0$ are real numbers, μ_s is a Poisson random measure and $\tilde{\mu}_t = \mu_t - \nu_t$ is a compensated measure, and ν_t is a Lévy measure satisfying the usual condition. $(W_t)_{t\geq 0}$ is a Brownian motion that is independent of μ_t. The third term is a compound Poisson process and the last term is a pure jump process with an infinite number of jumps.

4.4 Model assumptions

The Lévy instantaneous forward rate

$$f(t,T) = f(0,T) + \alpha_{\text{Lévy}}(t,T)dt - \sigma(t,T)dL_t \;\; \forall 0 \leq t \leq T \leq \mathbb{T},$$

with initial values $f(0,T)$ which is assumed to be deterministic, measurable, and bounded. The following conditions by Eberlein and Kluge (2006) are imposed on $\alpha_{\text{Lévy}}$ and σ.

Condition 4.4.1 *(C1) $(\omega, s, T) \mapsto \alpha_{Lévy}(\omega, s, T)$ and $(\omega, s, T) \mapsto \sigma(\omega, s, T)$ are measurable with respect to $\mathbb{P} \otimes \mathcal{B}([0, \mathbb{T}])$.*

(C2) For $s > T$, the following holds:

$$\alpha_{Lévy}(\omega, s, T) = 0 \quad \textit{and} \quad \sigma(\omega, s, T) = 0, \;\; \forall\, \omega.$$

(C3)

$$\sup_{s,T\leq\mathbb{T}} \left(|\alpha_{Lévy}(\omega, s, T)| + ||\sigma(\omega, s, T)||\right) < \infty, \;\; \forall\, \omega.$$

We have shown that bond price dynamics have the form shown in Equation 4.3.5,

$$P(t,T) = P(0,T)B_t \exp\left(-\int_0^t \psi_s\left(S(s,T)\right)\, ds + \int_0^t S(s,T)\, dL_s\right).$$

To express it in an exponential form, let $t = T$ and substitute in the bond prices formula above, to get

$$1 = P(t,t) = P(0,t)B_t \exp\left(-\int_0^t \psi_s\left(S(s,t)\right)\, ds + \int_0^t S(s,t)\, dL_s\right).$$

Making the bank account B_t the subject,

$$B_t = \frac{1}{P(0,t)} \exp\left(\int_0^t \psi_s\left(S(s,t)\right)\, ds - \int_0^t S(s,t)\, dL_s\right). \tag{4.4.1}$$

The zero-coupon bond prices dynamics under Lévy process are given by

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp\left(\int_0^t \psi_t\left(S(s,t)\right) - \psi_s\left(S(s,T)\right)\, ds + \int_0^t S(s,T) - S(s,t)\, dL_s\right). \tag{4.4.2}$$

Stellenbosch University https://scholar.sun.ac.za

We can split Equation 4.4.2 into a deterministic part and a stochastic part as follows:

$$P(t,T) = D(t,T)\exp\left(X_t\right), \; D(t,T) = \frac{P(0,T)}{P(0,t)}\exp\left(\int_0^t \psi_s\left(S(s,t)\right) - \psi_s\left(S(s,T)\right)\, ds\right),$$
$$X_t = \int_0^t S(s,T) - S(s,t)\, dL_s. \quad (4.4.3)$$

Since L is a Lévy process generated by an infinitely divisible distribution and the volatility structure σ is assumed to be deterministic for all maturities, it turns out that the process X_t has independent increments although not necessary stationary (non-homogeneous). As mentioned by Kluge (2005), the process X_t is the underlying stochastic process to be investigated.

4.5 Volatility structure

Several assumptions are made before the implementation of Lévy HJM models. These assumptions include the nature of the volatility structure, which must be deterministic , bounded and twice differentiable with respect to both variables (refer to Condition 4.3.4). Based on Theorem 4.2.2, the choice of the volatility structure also has a significance, since it determines whether the short rate inferred from bond dynamics is Markovian, (see Carverhill, 1994) and Ritchken and Sankarasubramanian (1995). For Lévy HJM models, Eberlein and Raible (1999) proved that short rates derived from bond prices are Markov processes if the structure has the form for Vasiček or Ho-Lee volatility. We consider the volatility structure:

Condition 4.5.1 *For every $T \in [0, \mathbb{T}]$, we assume that $\sigma(\cdot, T)$ is non-zero and has a factorisation*

$$\sigma(t,T) = \sigma_1(t)\sigma_2(T), \;\; \textit{for all} \;\; (0 \le t \le T),$$

where $\sigma_1 : [0, \mathbb{T}] \mapsto (0, +\infty)$ and $\sigma_2 : [0, \mathbb{T}] \mapsto (0, +\infty)$ are continuously differentiable functions.

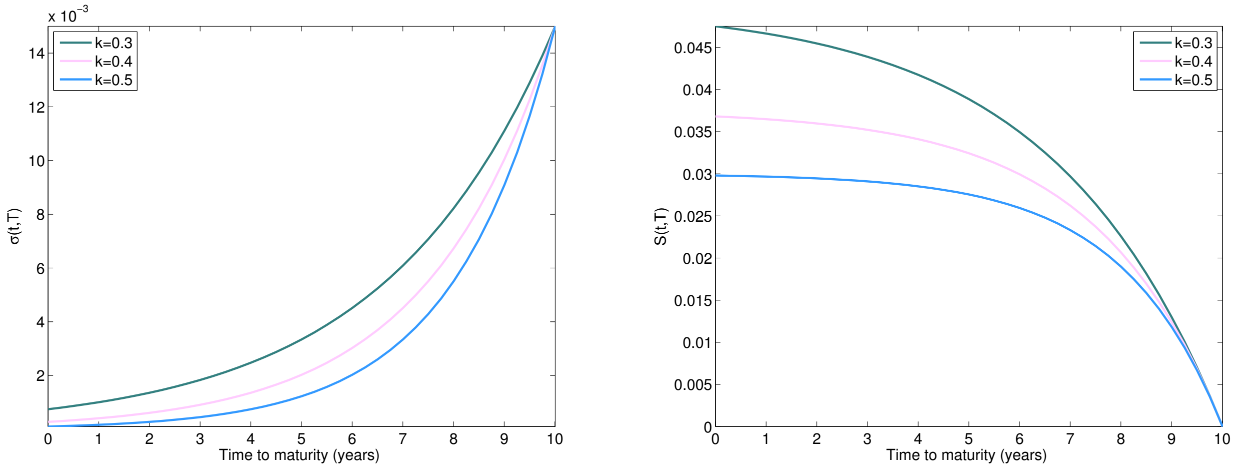
We take the volatility structure of the form for Vasiček, i.e.

$$\sigma(t,T) = \hat{\sigma} e^{-a(T-t)},$$

which means that the bond volatility is

$$S(t,T) = \frac{\hat{\sigma}}{a}\left(1 - e^{-a(T-t)}\right), \;\; \hat{\sigma} > 0, \; a \neq 0.$$

Stellenbosch University https://scholar.sun.ac.za



(a) Forward rate volatility (b) Bond volatility

Figure 4.1: Vasiček volatility, $\hat{\sigma} = 0.015$ and varying a.

Bond volatilities tend to zero as the bond approaches the time of maturity because of the effect of pull-to-par.

4.6 Options on zero-coupon bonds

In the previous section we derived a formula for zero-coupon bonds where the driver is a general Lévy process and is expressed in a general exponential Lévy form. It was also shown that bond prices, when deflated by the bank account, form a martingale, which is a crucial concept in option valuation theory. The task now is to compute values of options on zero-coupon bonds.

Suppose we want to value a put option with a strike price K and maturity time T on a zero-coupon bond maturing at a later time $U \geq T$. For a call option, this can be done in a similar way. The risk-neutral valuation principle in Definition 2.1.6 says that if $\Phi(P(T,U))$ represents some contingent claim that pays an unknown amount at time T, then the time $t = 0$ value is given by

$$\mathbb{V}_0 = B_0 \mathbb{E}_{\mathbb{Q}} \left[\frac{\Phi(P(T,U))}{B_T} \middle| \mathcal{F}_0 \right].$$

If Φ represents a zero-coupon bond maturing at time T, then the time t value is

$$\mathbb{V}_t = B_t \mathbb{E}_{\mathbb{Q}} \left[\frac{P(T,T)}{B_T} \right] = B_t \mathbb{E}_{\mathbb{Q}} \left[\frac{1}{B_T} \middle| \mathcal{F}_t \right].$$

Stellenbosch University https://scholar.sun.ac.za

The time t value of a put option $(\Phi(P(T,U)) = \max\{K - P(T,U), 0\})$ on a zero-coupon bond is given as

$$\mathbb{V}_{\text{put}}(t,T,U,K) = B_t \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B_T}\left(K - P(T,U)\right)^+ \middle| \mathcal{F}_t\right]. \tag{4.6.1}$$

$$\begin{aligned}
\frac{\mathbb{V}_{\text{put}}(t)}{B_t} = \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B_T}\left(K - P(T,U)\right)^+ \middle| \mathcal{F}_t\right] &= K\mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B_T}\mathbb{I}_{\{P(T,U)<K\}} \middle| \mathcal{F}_t\right] - \mathbb{E}_{\mathbb{Q}}\left[\frac{P(T,U)}{B_T}\mathbb{I}_{\{P(T,U)<K\}} \middle| \mathcal{F}_t\right] \\
&= K\mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B_T}\mathbb{I}_{\{P(T,U)<K\}} \middle| \mathcal{F}_t\right] - \mathbb{E}_{\mathbb{Q}}\left[\frac{B_T \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B_U} \middle| \mathcal{F}_T\right]}{B_T}\mathbb{I}_{\{P(T,U)<K\}} \middle| \mathcal{F}_t\right] \\
&= K\mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B_T}\mathbb{I}_{\{P(T,U)<K\}} \middle| \mathcal{F}_t\right] - \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B_U}\mathbb{I}_{\{P(T,U)<K\}} \middle| \mathcal{F}_t\right]
\end{aligned} \tag{4.6.2}$$

In computational terms, to compute these expectations, firstly one needs to know the joint density distribution function of the final bank account B_T, B_U and zero-coupon bond prices $P(T,U)$ because they may be correlated, i.e. we must know the dependence between B_T, B_U and $P(T,U)$. Although the joint density function has been found by Eberlein and Raible (1999), it is quite intensive in terms of computation as it takes time and is very cumbersome. At first glance, the ideal tool for computing the expectations in Equation 4.6.1 is the Monte Carlo simulation.

4.7 The Monte Carlo method

This section illustrates how to solve the expectations in Equation 4.6.1 using the Monte Carlo method. Monte Carlo is a powerful numerical method that is widely used as a benchmark method in computational finance. Its main advantage is that it is straightforward and extremely easy to implement. We consider an example of a put option on a zero-coupon bond in a general Lévy HJM framework.

Recall that the time t value of a put option on a zero-coupon bond under the equivalent martingale measure $\mathbb{Q}$ is given by

$$\mathbb{V}_{\text{put}}(t,T,U,K) = B_t \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B_T}\max\{0, K - P(T,U)\} \middle| \mathcal{F}_t\right].$$

Define a process

$$Z(t,T,u) = \exp\left(-\int_t^T \psi_s(S(s,u))\,ds + \int_t^T S(s,u)\,dL_s\right) = Z_u,$$

where the driver is the process

$$Y_T = \int_t^T S(s,T)\,dL_s.$$

Stellenbosch University https://scholar.sun.ac.za

Recall from Equation 4.4.1 that the discount factor $\frac{1}{B_T}$ is given by

$$\frac{1}{B_t} = P(0,T)\exp\left(-\int_0^T \psi_s\left(S(s,T)\right)\, ds + \int_0^T S(s,T)\, dL_s\right) = P(0,T)Z(t,T,T)$$

and the expression for the quotient $\frac{P(T,U)}{B_T}$ is given by

$$\frac{P(T,U)}{B_T} = P(0,U)\exp\left(-\int_t^T \psi_s\left(S(s,U)\right)\, ds + \int_t^T S(s,U)\, dL_s\right) = P(0,U)Z(t,T,U).$$

Hence, the time t value of a put option on a zero-coupon bond can be rewritten as

$$\mathbb{V}_{\text{put}}(t,T,U,K) = B_t\mathbb{E}_{\mathbb{Q}}\left[\max\{0, KP(t,T)Z(t,T,T) - P(T,U)Z(t,T,U)\}\right].$$

At time $t = 0$, we can now use Monte Carlo techniques by simulating random variables Z_T and Z_U as follows:

Algorithm *M*onte Carlo simulation$(P(0,U), P(0,T), U, T, M)$

1. **for** $i \leftarrow 1$ **to** N
2. $t \leftarrow t_0, \cdots, t_N = T$
3. **do** $\Delta t = \frac{T}{N}$
4. $t_{i+1} \leftarrow t_i + \Delta t$
5. Generate a process L and $\Delta L_i \leftarrow L_i - L_{i-1}$
6. $Z_i(0,T,s) \leftarrow \exp\left(\sum_{i=1}^{N}\left(-\psi(S(t_i,s))\,\Delta t + S(t_i,s)\,\Delta L_{u_j}\right)\right)$
7. $$\mathbb{V}_{\text{put}}(0) = \lim_{M\to\infty}\frac{1}{M}\sum_{i=1}^{M}\max\{0, KP(0,T)Z_i(0,T,T) - P(T,U)Z_i(0,T,U)\}.$$

Furthermore, the expectation in Equation 4.8.4 can be solved provided the joint density function of the random variables

$$Y_T = \int_0^T S(s,T)\, dL_s \quad \text{and} \quad Y_U = \int_0^T S(s,U)\, dL_s$$

is known. Eberlein and Raible (1999) showed that the joint distribution of these random variables possesses a Lebesgue density and found the analytical expression for the corresponding characteristic function based on Theorem 4.3.2:

$$\mathbb{E}_{\mathbb{Q}}[\exp(\mathrm{i}uY_T + \mathrm{i}vY_U)] = \exp\left(\int_0^T \psi_s(uS(s,U) + vS(s,T))\, ds\right), \tag{4.7.1}$$

in which the density for both X_U and X_T can be recovered using two-dimensional FFT.

Stellenbosch University https://scholar.sun.ac.za

4.8 Change of numéraire

In this section we explore the methodology of changing the numéraire as a way of simplifying the computations of the expectation in Equation 4.6.1.

Recall that a numéraire is any positively priced asset that does not pay dividends. For every numéraire there is an associated probability measure. Since zero-coupon bonds are assets, we will consider changing the numéraire from bank account B_t to a zero-coupon bond whose maturity coincides with the maturity time T of the option. The associated measure is called T-forward martingale measure, denoted by $\mathbb{Q}^T$, i.e. an equivalent martingale measure defined with respect to the bond forward prices. The reader interested in the changes of numéraire is referred to Geman *et al.* (1995).

Definition 4.8.1 (T-forward measure) *Keep the maturity time T fixed. The T forward measure $\mathbb{Q}^T$ is defined by*

$$\mathbb{Q}^T(A) = \frac{1}{P(0,T)} \int_A Z(\omega)\, d\mathbb{Q}(\omega) \ \textit{for all} \ \ A \in \mathcal{F},$$

where the Radon–Nikodym derivative process $Z = Z_T$ is given by

$$Z_T = \frac{d\mathbb{Q}^T}{d\mathbb{Q}} = \frac{1}{B_T P(0,T)} = \exp\left(-\int_0^T \psi_s\left(S(s,T)\right)\, ds + \int_0^T S(s,T)\, dL_s\right)$$

and when conditioned on $\mathcal{F}_t$,

$$Z_t = \mathbb{E}_{\mathbb{Q}}\left[\frac{P(t,T)B_0}{B_T P(0,T)}\middle|\mathcal{F}_t\right] = \frac{B_t^{-1}P(t,T)}{P(0,T)} = \exp\left(-\int_0^t \psi_s\left(S(s,T)\right)\, ds + \int_0^t S(s,T)\, dL_s\right).$$

The Radon–Nikodym derivative above can be interpreted in two trades:

(a) Denominator $P(0,T)$: looks as if at time $t=0$ we bought a T-bond which cost us $P(0,T)$ and pays 1 at time $t=T$.

(b) Numerator $\frac{P(t,T)}{B_t}$: Reflects that at time t we purchased a zero-coupon bond maturing at time T, costing us $P(t,T)$, and still paying 1 at time $t=T$. This means that the time $t=0$ the value is $\frac{P(t,T)}{B_t}$.

Both the numerator and denominator end up paying 1 at maturity time $t=T$, so to avoid arbitrages opportunities we must have an expectation of 1. To make sure that the chosen Radon–Nikodym derivative is correct, one can show that it is a martingale by showing that $\mathbb{E}\left[Z_T|\mathcal{F}_t\right] = 1$. This is a one-line proof because we know that $\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\int_0^t S(s,T)\, dL_s\right)\right] = \exp\left(\int_0^t \psi_s\left(S(s,T)\right)\, ds\right)$ based on Theorem 4.3.2 and both $\psi_s(u)$ and $S(t,T)$ are deterministic

Stellenbosch University https://scholar.sun.ac.za

functions for all maturities. Therefore we have

$$\begin{aligned}\mathbb{E}\left[Z_T|\mathcal{F}_t\right] &= \mathbb{E}\left[\exp\left(-\int_0^T \psi_s\left(S(s,T)\right)\,ds + \int_0^T S(s,T)\,dL_s\right)\Big|\mathcal{F}_t\right] \\ &= \exp\left(-\int_0^T \psi_s\left(S(s,T)\,ds\right)\right)\mathbb{E}\left[\exp\left(\int_0^T S(s,T)\,dL_s\right)\Big|\mathcal{F}_t\right] \\ &= \exp\left(-\int_0^t \psi_s\left(S(s,T)\right)\,ds\right)\exp\left(\int_0^t \psi_s\left(S(s,T)\right)\,ds\right) = 1.\end{aligned} \tag{4.8.1}$$

The Radon–Nikodym derivative is correct as long as the volatility structure σ is finite and deterministic. Ultimately, this means Z_t is an exponential martingale. Consequently, the forward measure $\mathbb{Q}^T$ is well-defined.

Under $\mathbb{Q}^T$, it follows from the Girsanov theorem for Lévy process (see Theorem 3.5.4) that the non-homogeneous Lévy process has the following Lévy triplet:

$$\left(a_s^{\mathbb{Q}^T}, b_s^{\mathbb{Q}^T}, \nu_s^{\mathbb{Q}^T}\right),$$

defined by

$$\begin{aligned}a_s^{\mathbb{Q}^T} &= a_s + b_s S(s,T) + \int_{\mathbb{R}} \left(e^{S(s,T)x} - 1\right) x\,\nu_s(dx), \\ b_s^{\mathbb{Q}^T} &= b_s, \\ \nu_s^{\mathbb{Q}^T} &= e^{S(s,T)x}\nu_s(dx),\end{aligned} \tag{4.8.2}$$

and that the $\mathbb{Q}^T$ Brownian motion is given by

$$W_t^{\mathbb{Q}^T} = W_t^{\mathbb{Q}} - \int_0^t \sqrt{b_s} S(s,T)\,ds.$$

Observe that under the forward martingale measure $\mathbb{Q}^T$, the Lévy triplets $\left(a_s^{\mathbb{Q}^T}, b_s^{\mathbb{Q}^T}, \nu_s^{\mathbb{Q}^T}\right)$ are deterministic functions of time and hence L_t has deterministic Lévy triplets which are absolutely continuous. L_t retains its non-homogeneity properties. This is one of the important fundamental properties that motivates the use of non-homogeneous Lévy processes. For instance, if one chooses to apply a homogeneous Lévy process to a driving process then after the change of measure, then the driving process is no longer a homogeneous Lévy process. Let the pay-off function be defined by

$$\Phi : \mathbb{R} \mapsto [0,\infty) \text{ defined by } \Phi(P(T,U)) = \max\{K - P(t,U), 0\}. \tag{4.8.3}$$

Equation 4.6.1 becomes

$$\mathbb{V}_{\text{put}}\left(t,T,U,K\right) = B_t\mathbb{E}_{\mathbb{Q}}\left[\frac{\Phi(P(T,U))}{B_T}\Big|\mathcal{F}_t\right].$$

Stellenbosch University https://scholar.sun.ac.za

Divide both sides by $P(t,T)$, we get

$$\frac{\mathbb{V}_{\text{put}}(t,T,U,K)}{P(t,T)} = \frac{B_t}{P(t,T)}\mathbb{E}_{\mathbb{Q}}\left[\frac{\Phi(P(T,U))}{B_T}\bigg|\mathcal{F}_t\right] = \mathbb{E}_{\mathbb{Q}^T}\left[\Phi(P(T,U))|\mathcal{F}_t\right].$$

This gives a simplified version of the value of a put option on a zero-coupon bond, in which we are only required to compute the expectation of $f(T)$ instead of $\frac{\Phi(P(T,U))}{B_T}$.

$$\mathbb{V}_{\text{put}}(t,T,U,K) = P(t,T)\mathbb{E}_{\mathbb{Q}^T}\left[\Phi(P(T,U))|\mathcal{F}_t\right] \tag{4.8.4}$$

4.9 The Gaussian HJM

In this section we outline pricing of derivatives on zero-coupon bonds in the case where the driving process is a Brownian motion, i.e. $L_t = W_t$. Since we are working with deterministic volatility, this case is often referred to as the Gaussian HJM. This case is worth discussing due to the availability of analytical option price formulae. The main problem is evaluating the pricing problem in Equation 4.8.4.

Under the forward probability measure $\mathbb{Q}^T$, the forward zero-coupon bond prices in Equation 2.2.3 seen at time $t=0$ for future settlements is calculated as follows:

$$\mathcal{B}(0,T,U) = \mathbb{E}_{\mathbb{Q}^T}\left[P(T,U)|\mathcal{F}\right] = \mathbb{E}_{\mathbb{Q}}\left[\frac{P(T,U)}{B_T P(0,T)}\bigg|\mathcal{F}_0\right] = \frac{1}{P(0,T)}\underbrace{\mathbb{E}_{\mathbb{Q}}\left[\frac{P(T,U)}{B_T}\bigg|\mathcal{F}_0\right]}_{P(0,U)}$$
$$= \frac{P(0,U)}{P(0,T)}.$$

Definition 4.9.1 (Forward bond price) *We call $\mathcal{B}(t,T,U)$ the forward bond price at time t for future settlement time T of a zero-coupon bond with maturity time $U > T$. It is defined by*

$$\mathcal{B}(t,T,U) := \frac{P(t,U)}{P(t,T)}. \tag{4.9.1}$$

The log moment generating function of a Brownian motion is given by

$$\psi(u) = \log\mathbb{E}\left[\exp(uW_1)\right] = \frac{1}{2}u^2. \tag{4.9.2}$$

The bond price process driven by a Brownian motion is therefore given by

$$P(t,T) = \frac{P(0,T)}{P(0,t)}\exp\left(\int_0^t \psi_s\left(S(s,t)\right) - \psi_s\left(S(s,T)\right)\, ds + \int_0^t S(s,T) - S(s,t)\, dW_s\right)$$
$$= \frac{P(0,T)}{P(0,t)}\exp\left(\frac{1}{2}\int_0^t ||S(s,t)||^2 - ||S(s,T)||\,|^2\, ds + \int_0^t S(s,T) - S(s,t)\, dW_s^{\mathbb{Q}}\right).$$

Stellenbosch University https://scholar.sun.ac.za

To derive the formula for $\mathcal{B}(t,T,U)$, we just have to know its volatility, which is the difference between volatility of $P(t,U)$ and $P(t,T)$. Therefore, the forward bond price is given by

$$\mathcal{B}(t,T,U) = \mathcal{B}(0,T,U)\exp\left(-\frac{1}{2}\int_0^t ||S(s,U)||^2 - ||S(s,T)||^2\, ds + \int_0^t S(s,U) - S(s,T)\, dW_s^{\mathbb{Q}}\right). \tag{4.9.3}$$

Rewrite Equation 4.9.3 as follows:

$$\mathcal{B}(t,T,U) = \mathcal{B}(0,T,U)\exp\left(-\frac{1}{2}\int_0^t S(s,T)\left[S(s,U) - S(s,T)\right]\, ds\right)\times \tag{4.9.4}$$
$$\exp\left(-\frac{1}{2}\int_0^t \left(S(s,U) - S(s,T)\right)^2\, ds + \int_0^t S(s,U) - S(s,T)\, dW_s^{\mathbb{Q}}\right).$$

In stochastic differential form

$$d\mathcal{B}(t,T,U) = \mathcal{B}(0,T,U)\left[-\frac{1}{2}S(t,T)\left[S(s,U) - S(t,T)\right]\, dt + S(t,U) - S(t,T)\, dW_t^{\mathbb{Q}}\right]. \tag{4.9.5}$$

Now we change the Brownian motion $W_t^{\mathbb{Q}}$ to a T-forward associated Brownian motion.This is done via the Girsanov transformation for Brownian motion (Theorem 3.5.5).

$$W_t^{\mathbb{Q}^T} = W_t^{\mathbb{Q}} - \int_0^T S(s,T)\, ds,$$

and the $\mathbb{Q}^T$ dynamics of the forward bond prices in differential form becomes

$$d\mathcal{B}(t,T,U) = \mathcal{B}(0,T,U)\left[S(t,U) - S(t,T)\right]\, dW_t^{\mathbb{Q}^T}. \tag{4.9.6}$$

This means that under the forward martingale measure $\mathbb{Q}^T$ the forward bond prices are martingales.

Consider a function $Y_t = \log \mathcal{B}(t,T,U)$. Using Itô's formula for stochastic differential equation,

$$dY_t = \frac{1}{\mathcal{B}_t}d\mathcal{B}_t - \frac{1}{2\mathcal{B}_t^2}\left(d\mathcal{B}_t\right)^2 = \left[S(t,U) - S(t,T)\right]\, dW_t^{\mathbb{Q}^T} - \frac{1}{2}||S(t,U) - S(t,T)||^2\, dt,$$

where we have used $(dW_t)^2 = dt$. This implies that the function $Y_t = \log \mathcal{B}(t,T,U)$ follows a Brownian motion with drift.

The solution is given by

$$Y_t = Y_0 - \frac{1}{2}||S(t,U) - S(t,T)||^2\, dt + \left[S(t,U) - S(t,T)\right]\, dW_t^{\mathbb{Q}^T}.$$

$$\mathcal{B}(t,T,U) = \mathcal{B}(0,T,U)\exp\left(-\frac{1}{2}\int_0^t ||S(s,U) - S(s,T)||^2\, ds + \int_0^t S(s,U) - S(s,T)\, dW_s^{\mathbb{Q}^T}\right).$$

We conclude that

$$Y_T \sim N\left(Y_0 - \frac{1}{2}||S(t,U) - S(t,T)||^2 T, \int_0^T ||S(s,U) - S(s,T)||^2\, dt\right).$$

Stellenbosch University https://scholar.sun.ac.za

In this a case zero-coupon bond prices are log-normally distributed.

We now solving the pricing problem under $\mathbb{Q}^T$:

$$\mathbb{V}_{\text{put}}(T,U,K) = P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[(K-P(T,U))^+\middle|\mathcal{F}_0\right] = P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[(K-\mathcal{B}(T,T,U))^+\middle|\mathcal{F}_0\right].$$

Consider the following important fact that will help in solving this expectation (see Musiela and Rutkowski, 2004, Chapter 11).

Theorem 4.9.2 *Assume that Y_t is log-normally distributed under $\mathbb{P}$, i.e. $Y_t \sim N(\mu, v^2)$. For any positive constant K, we have*

$$\mathbb{P}(Y \geq K) = N(-d_-) \;\; \textit{and} \;\; \mathbb{E}_{\mathbb{P}}\left[(K-Y)^+\right] = KN(-d_2) - \mathbb{E}_{\mathbb{Q}^T}[Y]\, N(-d_1),$$

where

$$d_\pm = \frac{\log\frac{\mathbb{E}_{\mathbb{P}}[Y]}{K} \pm \frac{1}{2}v^2}{v}$$

and

$$N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{\frac{-t^2}{2}}\, dt.$$

We apply Theorem 4.9.2 as follows:

$$Y_t = \log\mathcal{B}(t,T,U) \Longrightarrow \mathbb{E}_{\mathbb{Q}^T}[\mathcal{B}] = \mathcal{B}(0,T,U) = \frac{P(0,U)}{P(0,T)} \;\text{ and }\; v^2 = \int_0^T ||S(s,U)-S(s,T)||^2\, dt.$$

Since the forward bond prices are martingales and their instantaneous volatilities are deterministic, we have the following formulae

$$\begin{aligned}\mathbb{V}_{\text{put}}(0) &= P(0,T)\left(KN(-d_2) - \mathbb{E}_{\mathbb{Q}^T}[Y]\, N(-d_1)\right)\\ &= KP(0,T)N(-d_-) - P(0,U)N(-d_+)\end{aligned} \tag{4.9.7}$$

and

$$\begin{aligned}\mathbb{V}_{\text{call}}(0) &= P(0,T)\left(\mathbb{E}_{\mathbb{Q}^T}[Y]\, N(d_+) - KN(d_-)\right)\\ &= P(0,U)N(d_+) - KP(0,T)N(d_-)\end{aligned} \tag{4.9.8}$$

where $N(x)$ is the probability distribution function of a standard normal random variate.

Stellenbosch University https://scholar.sun.ac.za

Chapter 5

Valuation methods in the Lévy HJM framework

In this chapter, we discuss transform methods that play an increasingly important role in Mathematics of Finance. To this end, we compare different pricing methods in terms of speed and accuracy.

Recall from Equation 4.4.3 that a process $P(t,T) = D(t,T)\exp(X_t)$ is an exponential semimartingale process. We model a process

$$X_t = \int_0^t S(s,U) - S(s,T)\, dL_s$$

where $S(t,T)$ is the volatility function at time t with maturity time T. We consider a pay-off function

$$\Phi(x) = (K - e^x)^+. \tag{5.0.1}$$

Let $\mathcal{T}$ and $\mathcal{T}^{-1}$ be a transform method and its inverse respectively. By using an appropriate transform method, option valuation involves the following transformation steps (when the computation of $\mathcal{T}(\Phi(x))$ and $\mathcal{T}(d\mathbb{P}_X(x))$ is well defined):

$$\mathbb{E}_{\mathbb{Q}}\left[\Phi(X)\right] = \int_{\mathbb{R}} \Phi(x) d\mathbb{P}_X(x) \xrightarrow{\mathcal{T}} \mathcal{T}\left(\mathbb{E}_{\mathbb{Q}}\left[\Phi(X)\right]\right) \xrightarrow{\mathcal{T}^{-1}} \mathbb{E}_{\mathbb{Q}}\left[\Phi(X)\right] = \int_{\mathbb{R}} \mathcal{T}(\Phi(x))\mathcal{T}(d\mathbb{P}_X(x)), \tag{5.0.2}$$

where $\mathbb{P}_X$ represents the law of a random variable X.

The commonly used transform functions $\mathcal{T}$ are:

(a) Laplace transform

(b) Fourier transform

We begin with a review of the Fourier method of pricing.

Stellenbosch University https://scholar.sun.ac.za

5.1 General Fourier transform

This chapter introduces Fourier methods to solve the option pricing problem. The use of Fourier methods allows for rapid evaluation of complicated option pricing formulae. These methods are applicable as long as the characteristic function of the logarithm of the underlying variable exists. They also involve the Fourier transform of the pay-off function.

The use of Fourier methods has been motivated by Carr and Madan (1999), Raible (2000) and Eberlein *et al.* (2010). We take a unique approach to the derivation of option pricing formulae; we use Parseval's Theorem to derive option pricing formulae, following the ideas of Lewis (2001). The main advantage of using Parseval's Theorem is that we can compute the expectation in Equation 5.0.2 without needing to find the distribution of a random variable X.

5.2 Definitions

This subsection provides the essential definitions and results related to Fourier transform.

Denote by $L^1(\mathbb{R})$, a space of all functions $f : \mathbb{R} \mapsto \mathbb{C}$ with finite L^1-norm, i.e., such

$$||f||_{L^1} = \int_{\mathbb{R}} |f(x)|\, dx < \infty,$$

and $L^1_{bc}(\mathbb{R})$, a space of all bounded and continuous functions in $L^1(\mathbb{R})$.

Definition 5.2.1 (Fourier transform) *The Fourier transform of a suitably integrable function $f : \mathbb{R} \mapsto \mathbb{C}$ (if $f \in L^1$) denoted by $\hat{f}(x)$ is defined by*

$$\hat{f}(u) = \mathcal{F}[f](u) = \int_{\mathbb{R}} e^{\mathrm{i}ux} f(x)\, dx \quad , u \in \mathbb{C}, \tag{5.2.1}$$

and the inverse Fourier transform of a function $g : \mathbb{R} \mapsto \mathbb{R}$ is defined by

$$g(x) = \mathcal{F}^{-1}[g](x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\mathrm{i}ux} g(u)\, du, \tag{5.2.2}$$

where $\mathrm{i} = \sqrt{-1}$ is the imaginary unit.

The integrals in Equations 5.2.1 and 5.2.2 will converge if $f, g \in L^1$.

Some basic results:

(a) If $\mathbb{P}_X(x)$ is the distribution of X, then the characteristic function $\chi(u)$ of f is defined as

$$\chi(u) := \hat{\mathbb{P}}_X(x) = \int_{\mathbb{R}} e^{\mathrm{i}ux} \mathbb{P}_X(dx) = \mathbb{E}\left[e^{\mathrm{i}uX}\right].$$

(b) For a real-valued function f, we have the following:

Stellenbosch University https://scholar.sun.ac.za

1.
$$\overline{\hat{f}(u)} = \hat{f}(-u),$$

2.
$$\exp(\mathrm{i}ux)f(x) = \cos(ux)f(x) + \mathrm{i}\sin(ux)f(x),$$

and
$$\mathcal{F}^{-1}[\hat{f}](x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ux}\hat{f}(u)\,du.$$

This follows from the fact that the integral of an odd function is 0.

Parseval's Theorem, well-known in probability theory, will be used to derive the value of a put option or call option on a zero-coupon bond, is stated here without proof.

Theorem 5.2.2 (Parseval's Theorem) *Let $\mathbb{P}_X$ be a distribution of a random variable X. Suppose g is a real-valued function such that*

(a) $g \in L^1$

(b) $g(x^+)$ *and* $g(x^-)$ *both exist and the integrals*
$$\int_0^{\epsilon} \frac{g(x+t) - g(x^+)}{t}\,dt, \quad \int_{-\epsilon}^{0} \frac{g(x-t) - g(x^-)}{t}\,dt$$
are finite for some $\epsilon > 0$.

(c) The function $y \mapsto \mathbb{E}[g(y+X)]$ *is continuous at* $y = 0$*, then*
$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x)\mathbb{P}_X(dx) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(-u)\hat{\mathbb{P}}_X(u)\,du.$$

5.3 Application

Now let $\xi = -\log D(t,T)$. The analysis begins with the logarithmic return on the zero-coupon bond market, which is equivalent to the process $X_T = \log\frac{P(t,T)}{D(t,T)}$. The random variable X_T has a probability measure $\mathbb{Q}$.

Denoting $\mathbb{P}_{X_T}$ and M_{X_T} to be the law and moment generating function of a random variable X_T, i.e.
$$M_{X_T} = \mathbb{E}_{\mathbb{Q}^T}\left[e^{uX_T}\right] = \chi_{X_T}(-\mathrm{i}u) \quad \forall u \in \mathbb{C}.$$

The pay-off function is given by
$$\max\{K - P(T,U), 0\} = \max\{K - D(T,U)\exp(X_T), 0\} = \max\{K - \exp(X_T - \xi), 0\} = \Phi(X_T - \xi).$$

Stellenbosch University https://scholar.sun.ac.za

Kluge (2005) proposed that to evaluate

$$\mathbb{V}_{\text{put}}(0,T,U,K) = P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[\Phi(X_T-\xi)|\mathcal{F}_0\right]$$

it is required that the distribution of X_T with respect to the forward martingale measure $\mathbb{Q}^T$ possesses a density function, i.e. it is absolutely continuous with respect to Lebesgue measure on the real line. Conditions to ensure the existence of the density function in our context are given by (Raible, 2000, Proposition 4.4).

An alternative approach is to find the characteristic function of X_T under the T-forward measure $\mathbb{Q}^T$ then we can recover the Lebesgue density function from the characteristic function using the following inversion theorem.

Theorem 5.3.1 *Let f be a density function whose characteristic function is ϕ_{X_T}. Then*

$$\int_x^\infty f(s)\ ds = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{1}{\mathrm{i}u}\Re\left(e^{-\mathrm{i}ux}\chi_{X_T}(u)\right)\ du.$$

Proof. (See Carr and Madan, 1999). □

Recovering the distribution function is expressed in terms of the integral. Taking derivatives with respect to x on both side of the equation in Theorem 5.3.1 yields

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}ux}\chi_{X_T}(u)\ du. \tag{5.3.1}$$

The T-forward measure characteristic function of the underlying process X_T is calculated following Lemma 3.6.3:

$$\begin{aligned}
&\chi_{X_T}(u) = \mathbb{E}_{\mathbb{Q}^T}\left[\exp(\mathrm{i}uX_T)\right] = \mathbb{E}_{\mathbb{Q}}\left[\exp(\mathrm{i}uX_T)Z_T\right]\\
&= \exp\left(-\int_0^T \psi_s\left(S(s,T)\right)\ ds\right)\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\int_0^T S(s,T)\ dL_s\right)\exp\left(\mathrm{i}u\int_0^T S(s,U)-S(s,T)\ dL_s\right)\right]\\
&= \exp\left(-\int_0^T \psi_s\left(S(s,T)\right)\ ds\right)\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\int_0^T \mathrm{i}uS(s,U)+(1-\mathrm{i}u)S(s,T)\ dL_s\right)\right]\\
&= \exp\left(-\int_0^T \psi_s\left(S(s,T)\right)\ ds\right)\exp\left(\int_0^T \psi_s\left(\mathrm{i}uS(s,U)+(1-\mathrm{i}u)S(s,T)\right)\ ds\right)\\
&= \exp\left(\int_0^T \psi_s\left(\mathrm{i}uS(s,U)+(1-\mathrm{i}u)S(s,T)\right)\ ds - \int_0^T \psi_s\left(S(s,T)\right)\ ds\right).
\end{aligned}$$

We have the following theorem.

Theorem 5.3.2 *Let $S:[0,T]\mapsto\mathbb{R}$ be a smooth and bounded function, and L_t be a Lévy process. The characteristic function of a process $X_T = \int_0^T S(s,U)-S(s,T)\ dL_s$ under a forward martingale measure $\mathbb{Q}^T$ is given by*

$$\chi_{X_T}(u) = \exp\left(\int_0^T \psi_s\left(\mathrm{i}uS(s,U)+(1-\mathrm{i}u)S(s,T)\right) - \psi_s\left(S(s,T)\right)ds\right). \tag{5.3.2}$$

Stellenbosch University https://scholar.sun.ac.za

An easiest approach is the application of Parseval's Theorem 5.2.2. At first glance, notice that a pay-off function in Equation 5.0.1 is generally not integrable (not in L^1). This makes it difficult to apply Parseval's Theorem. To avoid this, we dampen the pay-off function by some exponential factor. This is closely related to the concept of Fourier transformation.

A Fourier transform has great advantages over ordinal function transform techniques because we are able to find the image of functions which are unbounded as well. Following Carr and Madan (1999) and Raible (2000), for any $R \in \mathbb{R}$ such that $M_{X_T}(R) < \infty$, we can multiply the pay-off function in Equation (5.0.1) by a dampening factor e^{-Rx} to obtain a new pay-off function g, (-known as the dampened pay-off), defined by

$$g(x) = e^{-Rx}\Phi(x). \tag{5.3.3}$$

As mentioned before, the pay-off function in Equation 5.0.1 is not integrable. The idea behind dampening the pay-off is guaranteeing integrability. For instance, when recovering a density function from the characteristic function we get a factor of $\frac{1}{u}$ as seen in Theorem 5.3.1. This makes it difficult to apply numerical integration because of the singularity at zero which eventually affects the implementation the fast Fourier transform techniques (see Carr and Madan (1999)).

The Fourier transform of the function g in Equation 5.3.10 is given by

$$\hat{g}(u) = \int_{\mathbb{R}} e^{iux} g(x)\, dx = \int_{\mathbb{R}} e^{(iu-R)x}\Phi(x)\, dx \quad , u \in \mathbb{C}.$$

Suppose we want to find the transform of Φ. If we attempt using the ordinal transform method, it blows up to infinity. We apply the Fourier transform method as follows:

Let u be such that $\Re(u) \in (-\infty, 0)$ (for a put option case),

$$\begin{aligned}
\hat{\Phi}(u) = \int_{\mathbb{R}} e^{\mathrm{i}ux}(K - e^x)^+\, dx = \int_{\mathbb{R}} e^{\mathrm{i}ux}(e^{\log K} - e^x)^+\, dx &= e^{\log K}\int_{\mathbb{R}} e^{\mathrm{i}ux}(1 - e^{x-\log K})^+\, dx \\
&= K\int_{\log K}^{+\infty} e^{\mathrm{i}ux} dx - \int_{\log K}^{+\infty} e^{(1+\mathrm{i}u)x}\, dx \\
&= \frac{Ke^{\mathrm{i}u\log K}}{\mathrm{i}u} - \frac{e^{(1+\mathrm{i}u)\log K}}{1+\mathrm{i}u} = \frac{K^{1+\mathrm{i}u}}{\mathrm{i}u} - \frac{K^{1+\mathrm{i}u}}{1+\mathrm{i}u} \\
&= \frac{K^{1+\mathrm{i}u}}{\mathrm{i}u(1+\mathrm{i}u)} \quad \forall \Re(u) \in (-\infty, 0).
\end{aligned} \tag{5.3.4}$$

The limit above only exists under the condition that $\Re(u) \in (-\infty, 0)$, which means that the Fourier transform is behaving well within that domain.

Stellenbosch University https://scholar.sun.ac.za

From this, we can deduce the Fourier transform of the dampened pay-off function g:

$$\begin{aligned}\hat{g}(u) &= \int_{\mathbb{R}} e^{\mathrm{i}ux} g(x)\, dx = \int_{\mathbb{R}} e^{\mathrm{i}ux} e^{-Rx} \Phi(x)\, dx \\ &= \int_{\mathbb{R}} e^{\mathrm{i}(u+\mathrm{i}R)x} \Phi(x)\, dx = \frac{K^{1+\mathrm{i}u-R}}{(\mathrm{i}u-R)(1+\mathrm{i}u-R)}, \quad \text{for } R \in (\infty, 0).\end{aligned} \tag{5.3.5}$$

Hence we can write $\hat{g}(u)$ as

$$\hat{g}(u) = \hat{\Phi}(u + \mathrm{i}R). \tag{5.3.6}$$

A Fourier transform of a dampened density ($P = e^{Rx}\mathbb{P}$) is given by

$$\hat{P}(u) = \hat{\mathbb{P}}(u - \mathrm{i}R).$$

In order to derive the value of the option with any arbitrary pay-off function Φ, the following conditions are imposed on the dampened function g (see Eberlein *et al.*, 2010, Theorem 3.2):

Condition 5.3.3 *Suppose there exists $R \neq 0$ such that the following holds:*

(C1) The dampened pay-off function $g(x) = e^{-Rx}\Phi(x)$ is bounded and continuous in $L^1_{bc}(\mathbb{R})$, i.e. $\int_{\mathbb{R}} |g(x)|\, dx < \infty$.

(C2) $\mathbb{E}[e^{RX_T}] < \infty$.

(C3) The Fourier transform of the dampened pay-off function $\hat{g}$ belongs to $L^1(\mathbb{R})$.

Theorem 5.2.2 says that if $\mathbb{P}_X$ is a probability density of a random variable X and if both $g, \left(e^{Rx}\mathbb{P}_X\right) \in L^1$, then

$$\begin{aligned}\mathbb{E}[\Phi(X)] = \int_{\mathbb{R}} \Phi(x)\mathbb{P}_X(dx) &= \int_{\mathbb{R}} e^{-Rx}\Phi(x) e^{Rx}\mathbb{P}_X(dx) \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(-u)\hat{P}_X(u)\, du = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{\Phi}(\mathrm{i}R - u)\hat{\mathbb{P}}_{X_T}(u - \mathrm{i}R)\, du,\end{aligned} \tag{5.3.7}$$

where

$$P_X(dx) = e^{Rx}\mathbb{P}_X(dx).$$

Furthermore,

$$\begin{aligned}\mathbb{V}_{\text{put}}(0, T, U, K) &= P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[\Phi(X_T - \xi)^+ \middle| \mathcal{F}_0\right] \\ &= P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[e^{R(X_T-\xi)} g(X_T - \xi) \middle| \mathcal{F}_0\right] \\ &= P(0,T)e^{-R\xi}\int_{\mathbb{R}} e^{Rx} g(x - \xi)\mathbb{P}_X(dx) \\ &= \frac{P(0,T)}{2\pi} e^{-R\xi}\int_{\mathbb{R}} \hat{g}(-u - \xi)\hat{P}_X(u)\, du \\ &= \frac{P(0,T)}{2\pi} e^{-R\xi}\int_{\mathbb{R}} e^{-\mathrm{i}u\xi}\hat{g}(-u)\hat{P}_X(u)\, du.\end{aligned} \tag{5.3.8}$$

Stellenbosch University https://scholar.sun.ac.za

It remains to find

$$\begin{aligned}\hat{P}_X(u) = \mathbb{E}\left[e^{\mathrm{i}u+R}\right] &= \int_{\mathbb{R}} e^{\mathrm{i}ux}e^{Rx}P_X(x)\,dx \\ &= \int_{\mathbb{R}} e^{\mathrm{i}(u-\mathrm{i}R)x}\mathbb{P}_X(x)\,dx \\ &= \hat{\mathbb{P}}_X(u-\mathrm{i}R)\end{aligned} \tag{5.3.9}$$

It is well-known that the Fourier transform of a probability density function is the characteristic function, i.e. $\hat{\mathbb{P}}_X(u-\mathrm{i}R) = \chi_{X_T}(u-\mathrm{i}R)$.

Putting these pieces together, we have

$$\begin{aligned}\mathbb{V}_{\text{put}}(0,T,U,K) &= \frac{P(0,T)}{2\pi}e^{-R\xi}\int_{\mathbb{R}} e^{-\mathrm{i}u\xi}\hat{g}(-u)\chi_{X_T}(u-\mathrm{i}R)\,du \\ &= \frac{P(0,T)}{2\pi}e^{-R\xi}\int_{\mathbb{R}} e^{-\mathrm{i}u\xi}\frac{K^{1-\mathrm{i}u-R}}{(-\mathrm{i}u-R)(1-\mathrm{i}u-R)}\chi_{X_T}(u-\mathrm{i}R)\,du\end{aligned}$$

Since $\hat{g}(u) = \hat{\Phi}(u+\mathrm{i}R)$ based on Equation 5.3.6 it follows that

$$\hat{g}(-u) = \hat{\Phi}(\mathrm{i}R-u). \tag{5.3.10}$$

We summarise these in the following theorem, in which the option price is represented as a convolution of functions $\hat{\Phi}$ and χ_{X_T}.

Theorem 5.3.4 *Suppose the bond price follows a process $P(t,T) = De^{X_T}$. Let $\xi = -\log D$. For any pay-off function $\Phi(x)$ satisfying the conditions 5.3.3. Then the arbitrage price for a put option is given by*

$$\mathbb{V}_{put}(0,T,U,K) = \frac{P(0,T)}{2\pi}e^{-R\xi}\int_{\mathbb{R}} e^{-\mathrm{i}u\xi}\hat{\Phi}(\mathrm{i}R-u)\chi_{X_T}(u-iR)\,du, \tag{5.3.11}$$

where $R \in (-\infty, 0)$.

For a call option with a pay-off function $\Phi(x) = (e^x - K)^+, x = X_T - \xi$, has the same transform function $\hat{\Phi}$ as for a put option discussed above; therefore we can apply Theorem 5.3.4 for all $R \in (1,\infty)$. This theorem was first stated by Raible (2000), followed by Eberlein and Kluge (2006), who derived the result from the convolution representations that make use of the stochastic Fubini theorem. In this context we derived the result using the well-known theorem of Parseval.

An arsenal of numerical methods exists to compute the option value in Equation 5.3.11. Eberlein and Kluge (2006) derived explicit formulae for valuation based on Fourier transform methods which were further discussed by Eberlein *et al.* (2010) while Kuan and Webber (2001) proposed the use of a random trinomial lattice tree. The origin of all these numerical algorithms originated from (Raible, 2000, Section 3.3).

Stellenbosch University https://scholar.sun.ac.za

5.4 Option pricing using direct integration

The integral in Theorem 5.3.4, Equation 5.3.11 can be calculated using direct integration. In this section we show how this is done.

Let $h(u) = \hat{\Phi}(\mathrm{i}R - u)\chi_{X_T}(u - iR)$. The problem is to compute the integral

$$I(\xi) = \frac{e^{-R\xi}}{2\pi} \int_{\mathbb{R}} e^{-\mathrm{i}u\xi} h(u)\, du, \tag{5.4.1}$$

which is a real number since the Fourier transform of $h(u)$ is odd and even in its imaginary and real part respectively. Since we are computing option prices, it makes sense to treat $I(\xi)$ as an even function; therefore $I(\xi)$ becomes:

$$I(\xi) = \frac{e^{-R\xi}}{\pi} \int_0^{\infty} e^{-\mathrm{i}u\xi} h(u)\, du. \tag{5.4.2}$$

Let u_b to be an upper bound for integration. Then Equation 5.4.2 can be implemented using Simpson's quadrature rule. Let N be the number of equidistant intervals such that $du = \frac{du}{N}$ denotes the distance between integration points and $u_j = (j-1)du \;\; \forall j = 1, 2, \cdots N+1$ is the endpoints for integration interval. Then

$$\begin{aligned} I(\xi) &= \frac{e^{-R\xi}}{\pi} \int_0^{\infty} e^{-\mathrm{i}u\xi} h(u)\, du \approx \frac{e^{-R\xi}}{\pi} \int_0^{ub} e^{-\mathrm{i}u\xi} h(u)\, du \\ &\approx \frac{e^{-R\xi} du}{3\pi} \left(e^{-\mathrm{i}u_1\xi} h(u_1) + 4e^{-\mathrm{i}u_2\xi} h(u_2) + 2e^{-\mathrm{i}u_3\xi} h(u_3) + \cdots + 4e^{-\mathrm{i}u_N\xi} h(u_N) + e^{-\mathrm{i}u_{N+1}\xi} h(u_{N+1}) \right) \\ &\approx \frac{e^{-R\xi}}{\pi} \sum_{i=1}^{N+1} e^{-\mathrm{i}u_j\xi} h(u_j)\omega_j, \end{aligned} \tag{5.4.3}$$

where $\omega_j = \frac{du}{3}(3 + (-1)^j - \delta_{j-(N+1)})$ and

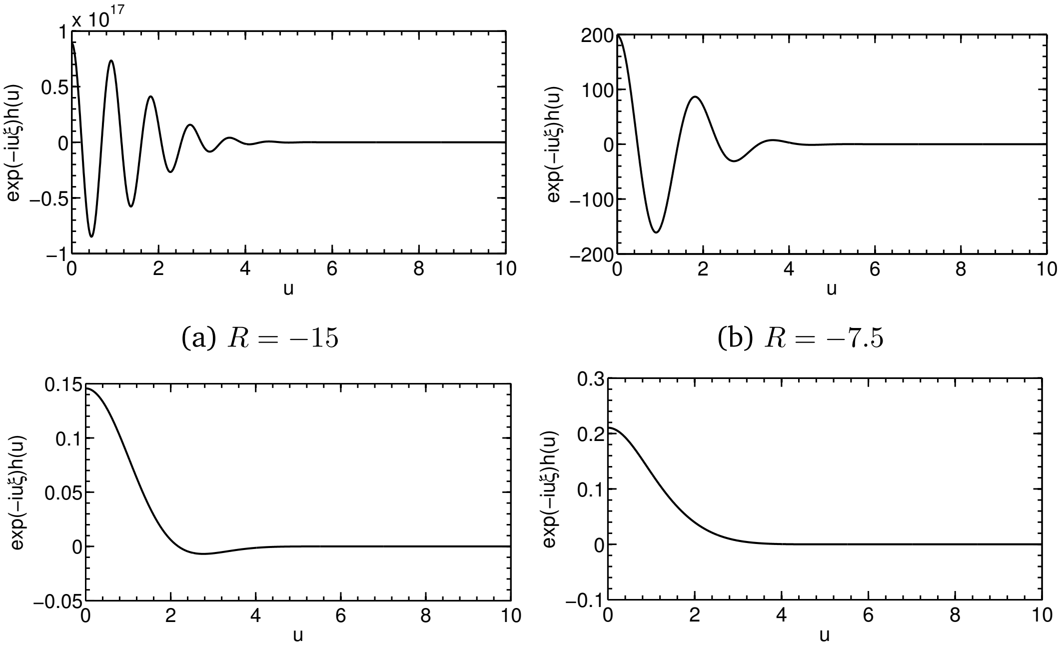
$$\delta_i = \begin{cases} 1, & \text{for j=0} \\ 0, & \text{otherwise.} \end{cases}$$

5.5 The role of a dampening parameter R

We have seen that a pay-off function $\Phi(x)$ is generally integrable. Hence we introduced the notion of dampening it to allow convergence. The choice of a dampening parameter R is very important as it affects the accuracy of the option value. A large R will cause the values of the characteristic function to approach zero which will eventually make it difficult to integrate accurately. On the other hand, if the value of R is close to zero, then we have the problem of singularity. One main task is to write algorithms for finding the optimal value for R. This is, however, is outside the scope of this study, but we shall demonstrate how sensitive the results

Stellenbosch University https://scholar.sun.ac.za

are due to the arbitrary choice of R. In other words, we examine how the choice of R affects the option value by plotting the integrand in Equation 5.4.2 and observing how fast the (tail of the) integrand $e^{-\mathrm{i}u\xi}h(u)$ converges to zero.

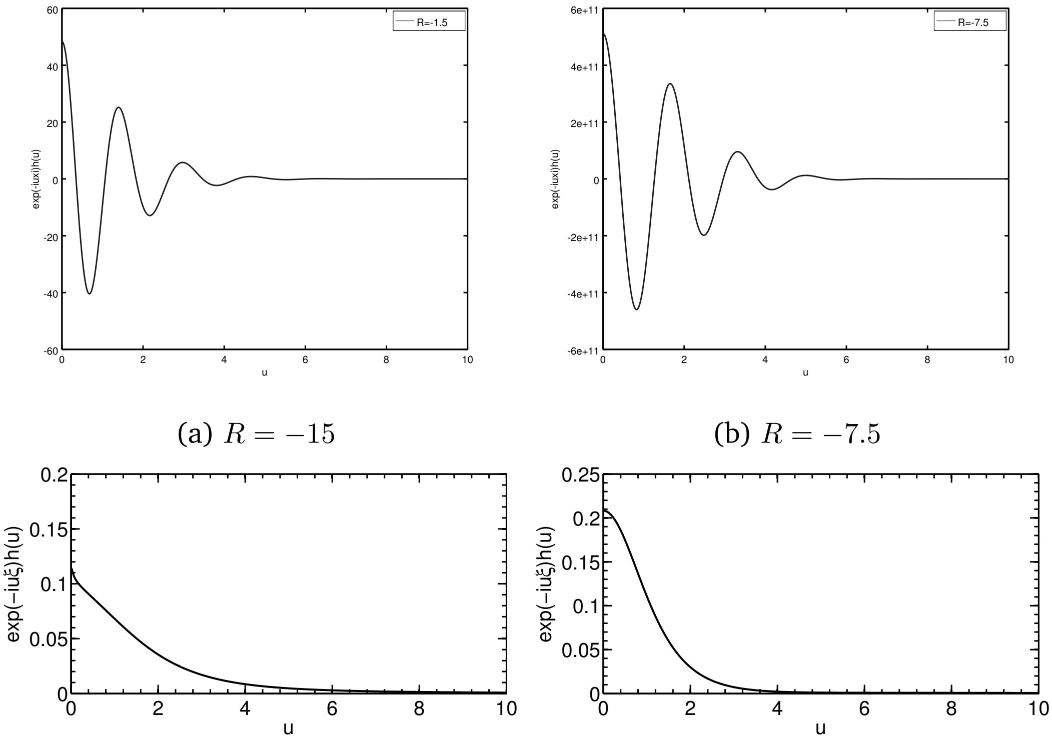


(a) $R = -15$ (b) $R = -7.5$

(c) $R = -2.5$ (d) $R = -1.5$

Figure 5.1: The effect of the choice of dampened parameter R for a model driven by a Brownian motion with parameters $\sigma = 1.5$ and $a = 0.5$.

Stellenbosch University https://scholar.sun.ac.za



(a) $R = -15$ (b) $R = -7.5$

(c) $R = -2.5$ (d) $R = -1.5$

Figure 5.2: The effect of the choice of dampened parameter R for a model driven by a generalised hyperbolic motion with parameters $\sigma = 1.5$, $a = 0.5$, $\lambda = 0.5$, $\alpha = 2$, $\beta = 0$, $\delta = 0.1$ and $\mu = 0$.

The above figures display the sensitivity of the integrand for the choice of R for a put option with strike $K = 0.9$, and maturity $T = 1$ on a zero-coupon bond maturing at time $T = 2$. As shown in both Figure 5.1 and 5.2, the greater the magnitude of R, the greater the oscillation of the integrand. In Figure 5.2, a high R makes the integrand decay exponentially, which results in poor approximation of the integral. Most stable values occur close to 0. The same analysis can be done for a call option, and one can verify that the most stable value of R is close to 1.

5.6 The Laplace method

This section reviews the numerical valuation based on the bilateral Laplace transform method introduced by Raible (2000).

The ideas of Raible (2000) are summarised below:

Stellenbosch University https://scholar.sun.ac.za

Raible's ideas:

$$\begin{aligned}\mathbb{V}_{\text{put}}(t) = P(t,T)\mathbb{E}_{\mathbb{Q}^T}[(K - P(T,U))^+] &= P(t,T)\int_\Omega (K - P(T,U))^+ \, d\mathbb{P} \\ &= P(t,T)\int_\Omega (K - De^x)^+ \, d\mathbb{P}(x) \\ &= P(0,T)\int_\mathbb{R} (K - De^x)^+ \mathbb{Q}^T(x).\end{aligned}$$

Now if $\mathbb{Q}^T$ is absolutely continuous with respect to the measure with density ρ, define $\Phi(x) = \max\{K - e^{-x}, 0\}$, and $\xi = -\log D(t,T)$, then

$$\mathbb{V} = P(t,T)\int_\mathbb{R} \Phi(\xi - x)\rho(x)\, dx = (\Phi * \rho)(\xi),$$

which is a convolution at point ξ. The Laplace transform of this convolution yields

$$\begin{aligned}\mathcal{L}_\mathbb{V}(u) = P(t,T)\int_\mathbb{R} e^{-ux}(\Phi * \rho)(x)\, dx &= P(t,T)\int_\mathbb{R} e^{-ux}\Phi(x)\, dx \int_\mathbb{R} e^{-ux}\rho(x)\, dx \\ &= P(t,T)\mathcal{L}_\Phi \mathcal{L}_\rho.\end{aligned}$$

$\mathcal{L}_\rho$ can be expressed in terms of the characteristic function, whereas $\mathcal{L}_\Phi$ can be calculated manually.

To find the option price, simply invert $\mathcal{L}_\mathbb{V}$. The final output can then be solved with the aid of the fast Fourier transform method.

5.7 Fast Fourier transform method

The integral in Equation 5.3.11 can be seen as an inverse Fourier transform of the function $H(x) = \hat{\Phi}(i\beta - u)\chi_{X_T}(u - i\beta)$. Hence one can apply fast Fourier transformation (FFT). FFT allows rapid calculation of the inverse Fourier transforms with vectors of many strikes very easily and efficiently.

The numerical procedure for FFT is outlined in Appendix A. To allow for ease of calculation, model parameters in the pay-off function can be reduced. This is done as follows:

We rewrite the pay-off function of a zero-coupon bond as a function of its deterministic part D, i.e. we let

$$\begin{aligned}&\mathbb{V}_{\text{put}}(0) = P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[(K - De^{X_T})^+ \middle| \mathcal{F}_0\right] \\ &= P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[(K - e^{-\xi}e^{X_T})^+ \middle| \mathcal{F}_0\right] \\ &= \mathbb{V}_{\text{put}}(0,T,U,K).\end{aligned}$$

Stellenbosch University https://scholar.sun.ac.za

Since $K = e^{\log K} > 0$, we proceed as follows

$$\begin{aligned}
\mathbb{V}_{\text{put}}(0,T,U,K) &= P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[(K - e^{-\xi}e^{X_T})^+ \middle| \mathcal{F}_0\right] \\
&= P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[(e^{\log K} - e^{-\xi}e^{X_T})^+ \middle| \mathcal{F}_0\right] \\
&= P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[e^{\log K}(1 - e^{-\xi-\log K}e^{X_T})^+ \middle| \mathcal{F}_0\right] \\
&= KP(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[(1 - e^{-(\xi+\log K)}e^{X_T})^+ \middle| \mathcal{F}_0\right] \\
&= KP(0,T)\mathbb{V}_{\text{put}}(0,T,U,1).
\end{aligned}$$

For $K = 1$, the Fourier transform of the function Φ becomes

$$\hat{\Phi}(u) = \frac{1}{\mathrm{i}u(1+\mathrm{i}u)}.$$

Now our pricing problem has reduced to

$$\mathbb{V}_{\text{put}}(0,T,U,K) = \frac{KP(0,T)}{2\pi}e^{-R\xi}\int_{\mathbb{R}} e^{-\mathrm{i}u\xi}\frac{1}{(-\mathrm{i}u-R)(1-\mathrm{i}u-R)}\chi_{X_T}(u-\mathrm{i}R)\,du. \tag{5.7.1}$$

The following algorithm is a modified version of Raible (2000). For notation refer to Appendix A.

Algorithm Application of FFT to bond pricing:

1. Choose $R \in \mathbb{R}$ such that $\chi_{X_T}(\mathrm{i}R) < \infty$ and $\hat{\Phi}(R) < \infty$.
2. **if** $R \in (-\infty, 0)$,
3. **then** (∗ Put option will be evaluated ∗),
4. **else** if $R \in (1, \infty)$,
5. **then** (∗ Call option will be evaluated ∗),
6. Choose step size Δu and $N = 2^m, \quad m \in \mathbb{R}^+$,
7. Compute the sequence $(y_n)_{n=1}^{N-1}$ defined by

$$y_n = \exp(-\mathrm{i}n\Delta u\gamma)\hat{\Phi}(\mathrm{i}R - n\Delta u)\chi_{X_T}(n\Delta u - \mathrm{i}R)\eta_n,$$

8. Apply FFT to y_n to get a sequences $(z_k)_{k=-\frac{N}{2}}^{\frac{N}{2}}$, i.e. $z_k = \text{FFT}(y_n)$
9. Compute D defined by

$$D = \frac{P(0,U)}{P(0,T)}\exp\left(\int_0^T \psi_s\left(S(s,T)\right) - \psi_s\left(S(s,U)\right)\,ds\right).$$

10. $\Delta x \leftarrow \frac{2\pi}{N\Delta u}$ so that $\forall k = -\frac{N}{2}, \cdots, \frac{N}{2}$, $\gamma = \frac{N\Delta x}{2}$, $x_k = -\gamma + k\Delta x = \log D$.
11. Compute $\log K_k = k\Delta x - x_k = \log D - x_k$,
12. For $K = 1$, the put option value is given by

$$\mathbb{V}_{\text{put}}(0,T,U,1) = \frac{P(0,T)\exp(-Rx_k)}{\pi}\Re\left\{z_k\right\}, \quad \text{where } x_k = -\gamma + k\Delta x.$$

Stellenbosch University https://scholar.sun.ac.za

13. For arbitrary K, the put option value is given by

$$\mathbb{V}_{\text{put}}(0,T,U,K_k) = K_k\mathbb{V}_{\text{put}}(x_k+\log K_k,1) = \frac{P(0,T)\exp(-\beta x_k)}{\pi}\Re\{z_k\}, \text{ where } x_k = \log D - \log K_k.$$

14. Interpolate prices using cubic splines.

5.8 Fractional fast Fourier transform

The fractional FFT (FrFT) attempts to solve a problem of the form

$$\int_0^\infty e^{-ixu}h(u)du \approx \sum_{j=0}^{N-1} e^{-i2\pi kj\alpha}h_j, \tag{5.8.1}$$

where α is the fractional parameter. It was introduced to finance by Chourdakis (2004).

The FFT is a special case of the FrFT, i.e. $\alpha = \frac{1}{N}$. Therefore the pricing formula is given by

$$\int_0^\infty e^{-ixu}h(u)du \approx \sum_{j=0}^{N-1} e^{-i2\pi\frac{1}{N}kj}h_j. \tag{5.8.2}$$

We can therefore calculate $\sum_{j=0}^{N-1} e^{-i\Delta x\Delta ukj}h_j$ without the restriction of $\Delta x\Delta u = \frac{2\pi}{N}$, which hampers the flexibility of the FFT method. This is because very small values of Δu may result in coarse output grids.

An N-point FrFT can be implemented by invoking two $2N$-points FFT and one inverse $2N$-points FFT steps and the selection of Δu and Δx is independent of each other.

Suppose we want to implement the FrFT procedures to compute $\vec{h} = (h_j)_{j=0}^{N-1}$. We then define two $2N$-dimensional vectors z_1 and z_2 by

$$z_1 = \left(\left(h_j e^{-i\pi j^2\alpha}\right)_{j=0}^{N-1}, (0)_{j=0}^{N-1}\right) \text{ and } z_2 = \left(\left(e^{i\pi j^2\alpha}\right)_{j=0}^{N-1}, \left(e^{i\pi(N-j)^2\alpha}\right)_{j=0}^{N-1}\right).$$

Then the discrete FrFT is given by

$$D_k(\vec{h},\alpha) = \left(e^{-i\pi j^2\alpha}\right)_{j=0}^{N-1} \odot D_k^{-1}\left(D_k(z_1)\odot D_k(z_2)\right), \tag{5.8.3}$$

where $\odot$ denotes component-wise multiplication.

Given the characteristic function of the underlying process X_t, choose Δu and Δx independently and set $\alpha = \frac{\Delta u\Delta x}{2\pi}$. The algorithm implements FrFT.

Algorithm Application of the FrFT method to bond pricing

1. Require parameter α, N, vector $\vec{x}$ and $\vec{u}$.
2. Choose Δu and Δx of your choice.

Stellenbosch University https://scholar.sun.ac.za

3. Compute z_1 and z_2.
4. Let $f_{z_1} = \text{FFT}(z_1)$ and $f_{z_2} = \text{FFT}(z_2)$.
5. Let $f_z = f_{z_1} \odot f_{z_2}$.
6. Let $if_z = \text{iFFT}(f_z)$.
7. Compute $f = \left(e^{-i\pi j^2 \alpha}\right)_{j=0}^{N-1} \odot if_z$.

The option price is obtained using the same approach as for the algorithm for application of FFT to bond pricing.

5.9 The COS method

Since we are interested in calibrating interest rate models to option prices, we seek an option pricing method which is sufficiently fast. We apply the Fourier cosine method (henceforth called the COS method) to bond option pricing. This method was introduced to finance by Fang and Oosterlee (2008) as an alternative method to the FFT and extended by Deng and Chong (2011) for options on stocks with many strikes. To our knowledge, the COS method has not yet been applied to bond option pricing in academic literature.

Back to our pricing problem for a European put option with maturity time T and strike price K on zero-coupon bond $P(\cdot, U)$ maturing at time $U \geq T$. We have seen from Equation 4.4.3 that bond prices can be reformulated in exponential Lévy form, i.e.

$$P(t,T) = P_t = D\exp(X_t) \iff X_t = \log\left[\frac{P_t}{D}\right].$$

We use the idea of Fang and Oosterlee (2008) to solve our problem. Let us assume Φ to be an even function. Then the Fourier transform based on the cosine transform is given by

$$\hat{\Phi}(u) = 2\int_0^{\infty} \Phi(x)\cos(xu)\,dx.$$

The theory of pricing European options with the cosine method originated from the fact that the density of log prices decays rapidly to zero as log prices tend to $\pm\infty$.

Recall that a Fourier series expansion of a function $\Phi(x) : [-\pi, \pi] \mapsto \mathbb{R}$ is the cosine series. Under some regularity conditions the function $\Phi : [0, \pi] \mapsto \mathbb{R}$ is also a cosine series.

The density function $\rho : [0, \pi] \mapsto \mathbb{R}$ can be represented by its cosine series expansion as follows:

$$\rho(x) = \frac{1}{2}A_0 + \sum_{n=1}^{\infty}\left[A_n\,\cos(n\,x)\right] \tag{5.9.1}$$

with Fourier cosine coefficient A_n

$$A_n = \frac{2}{\pi}\int_0^{\pi} f(y)\,\cos(n\,y)\,dy.$$

Stellenbosch University https://scholar.sun.ac.za

The Fourier series expansion for an even-density function $\rho(x)$ on an arbitrary finite interval $[a, b]$ can be found easily by changing variables that map $a \mapsto 0$ and $b \mapsto \pi$. We illustrate how this is done:

$$y = \frac{(x-a)\pi}{b-a}, \quad x = y + a.$$

Substitute this into Equation 5.9.1 to get

$$\rho(x) = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos\left(n\pi \frac{x-a}{b-a}\right) \quad \forall x \in [a, b], \tag{5.9.2}$$

where the Fourier cosine coefficients are given by

$$A_n = \frac{2}{b-a} \int_a^b \cos\left(n\pi \frac{x-a}{b-a}\right) \rho(x)\, dx, \quad n = 0, 1, 2, \cdots \tag{5.9.3}$$

Suppose that $\phi = \hat{\rho}$. It remains to express A_n of cosine expansion in terms of ϕ. We begin by truncating the characteristic function to a finite interval $[a, b]$,

$$\phi_1(u) = \int_a^b \exp(iux)\rho(x)\, dx + \epsilon \approx \phi(u), \tag{5.9.4}$$

where ϵ is the truncation error. We assume a and b are well chosen such that $\epsilon \to 0$. Recalling that $e^{iuy} = \cos(uy) + i\sin(uy)$ and comparing Equations 5.9.4 and 5.9.3, we get

$$A_n = \frac{2}{b-a} \Re\left\{\phi_1\left(\frac{n\pi}{b-a}\right) \exp\left(-i\frac{na\pi}{b-a}\right)\right\}, \quad n = 0, 1, 2, \cdots \tag{5.9.5}$$

For $n = 0, 1, \cdots, N-1$, we get an approximation for the density function

$$\rho(x) \approx \rho_1(x) = \frac{A_0}{2} + \sum_{n=1}^{N-1} A_n \cos\left(n\pi \frac{x-a}{b-a}\right) = \frac{A_0}{2} + \frac{2}{b-a} \Re\left\{\hat{\phi}_{X_T}\left(\frac{n\pi}{b-a}\right) \exp\left(-i\frac{na\pi}{b-a}\right)\right\} \tag{5.9.6}$$

$\forall x \in [a, b]$, where ρ_1 is known provided ϕ is known.

Suppose we would like to evaluate a put option, i.e., with a pay-off function $\Phi(P(T, U), K) = (K - De^{X_T})^+$. Then our pricing problem becomes

$$\begin{aligned}\mathbb{V}_{\text{put}}(0, T, U, K) &= KP(0, T) \int_0^{\infty} (1 - e^y)\rho(y)dy \approx KP(0, T) \int_0^b (1 - e^y) \left[\frac{A_0}{2} + \sum_{n=1}^{N-1} A_n \cos\left(n\pi \frac{y-a}{b-a}\right)\right] \\ &= KP(0, T) \left[\frac{A_0}{2}(\Psi_n(a, 0) - \Upsilon_n(a, 0)) + \sum_{n=1}^{N-1} A_n(\Psi_n(a, 0) - \Upsilon_n(a, 0))\right],\end{aligned}$$

where

$$\Upsilon_n(c, d) = \int_c^d \cos\left(n\pi \frac{x-a}{b-a}\right) dx \;\text{ and }\; \Psi_n(c, d) = \int_c^d e^x \cos\left(n\pi \frac{x-a}{b-a}\right) dx \quad \forall [c, d] \subset [a, b].$$

Stellenbosch University https://scholar.sun.ac.za

The analytical formulae for cosine series coefficients Υ_n and Ψ_n are given by direct integration and integration by parts respectively.

$$\Upsilon_n(c,d) = \int_c^d \cos\left(n\pi\frac{y-a}{b-a}\right)\,dy = \begin{cases} -\frac{b-a}{n\pi}\sin\left(n\pi\frac{y-a}{b-a}\right)\big|_c^d, & \text{for } n\neq 0 \\ d-c, & \text{for n=0.} \end{cases}$$

$$= \begin{cases} \frac{b-a}{n\pi}\left[\sin\left(n\pi\frac{c-a}{b-a}\right) - \sin\left(n\pi\frac{d-a}{b-a}\right)\right] & \text{for } n\neq 0 \\ d-c, & \text{for n=0.} \end{cases}$$

and

$$\Psi_n(c,d) = \int_c^d e^y \cos\left(n\pi\frac{y-a}{b-a}\right)\,dy = \frac{1}{1+\left(\frac{n\pi}{b-a}\right)^2}$$
$$\left[\cos\left(n\pi\frac{d-a}{b-a}\right)e^d - \cos\left(n\pi\frac{d-a}{b-a}\right)e^c + \frac{n\pi}{b-a}\sin\left(n\pi\frac{d-a}{b-a}\right)e^d - \frac{n\pi}{b-a}\sin\left(n\pi\frac{d-a}{b-a}\right)e^c\right].$$

The COS method does not need the use of a dampening parameter, which causes sensitivity to the option value and hence causes errors. However, the error within this method is due to the truncation of the integration range. To make sure that the truncation error is as small as possible, Fang and Oosterlee (2008) proposed that the range of the integration $[a,b]$ is chosen in the following manner:

$$[a,b] = \left[c_1 - L\sqrt{|c_2| + \sqrt{|c_4|}}, c_1 + L\sqrt{|c_2| + \sqrt{|c_4|}}\right], \; L = 10; \tag{5.9.8}$$

where L denotes truncation length and c_n denotes the n-cumulant of $\log\left(\frac{P(T,U)}{K}\right)$. It is also an open problem to find the optimal truncation. Moreover, Fang and Oosterlee (2008) illustrated that the COS method has a problem with short times to maturity. Furthermore, the approximation in Equation 5.9.8 is only accurate for $T \in [0.1, 10]$. Due to the complexity of our models, it is difficult to determine the cumulants, to circumvent this we propose the use of

$$[a,b] = [-2\sqrt{T}, 2\sqrt{T}] \tag{5.9.9}$$

which is a reliable choice for the size of the integration range. Our results proved to be stable for any maturity.

Algorithm Application of the cosine method to bond pricing

1. Compute $[a,b]$ as in (5.9.9)
2. Calculate

$$D_T = \frac{P(0,U)}{P(0,T)}\exp\left(\int_0^T \psi_s\left(S(s,T)\right) - \psi_s\left(S(s,U)\right)\,ds\right).$$

Stellenbosch University https://scholar.sun.ac.za

3. $x \leftarrow \log \frac{D_T}{K}$
4. **for** $n = 0, 1, \cdots, N-1,$
5. calculate $\Psi_n(x), \Upsilon_n(x), A_n(x)$.
6. Compute option value

$$\mathbb{V}_{\text{put}}(0,T,U,K) = KP(0,T)\left[\frac{A_0}{2}(\Psi_n(a,0) - \Upsilon_n(a,0)) + \sum_{n=1}^{N-1} A_n(\Psi_n(a,0) - \Upsilon_n(a,0))\right].$$

5.10 Numerical implementation

This section is devoted to the implementation of the FFT, the FrFT and the cosine method discussed above. We consider a call and put option on a zero-coupon bond and compare the pricing accuracy and speed between these methods. For a Gaussian case, we use the reference values from Equation 4.9.7, and for a model driven by a GH process we use the reference value from direct integration.

We have used a computer with the following specifications:

Computer specifications / features				
Processor number	**Catche**	**Clock speed**	**Memory type**	**Graphics**
Intel{®} Core{™} i7-5600U	4 MB	2.60 GHz	DDR3L–1333/1600	HD Graphics

We discuss Lévy HJM models driven by two processes, namely

(a) Brownian motion

(b) Generalised hyperbolic motion

Recall from Equation 5.3.2 that the characteristic function is given by

$$\chi_{X_T}(u) = \exp\left(\int_0^T \psi_t\left(\mathrm{i}uS(s,U) + (1-\mathrm{i}u)S(s,T)\right) - \psi_t\left(S(s,T)\right)ds\right).$$

Table 5.1: Log-moment generating function for the drivers.

Driver's name:	$\boldsymbol{\psi_t(u)} = \log\mathbb{E}\left[\mathbf{e^{uL_t}}\right]$
Brownian motion	$\frac{1}{2}u^2t$
Generalised hyperbolic (GH) motion	$\log\left[e^{\mu ut}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+u)^2}\right)^{\frac{\lambda t}{2}}\left(\frac{K_\lambda\delta\sqrt{\alpha^2-(\beta+u)^2}}{K_\lambda\delta\sqrt{\alpha^2-\beta^2}}\right)^t\right]$

Stellenbosch University https://scholar.sun.ac.za

We compute the root mean square error (RMSE) as follows:

$$\epsilon = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\{\mathbb{V}^i_{put}(0) - \hat{\mathbb{V}}^i_{put}(0)\right\}},$$

where $\hat{\mathbb{V}}_{put}(0)$ is the reference value. The dampening factor $R = 1.25$ for a call option and -1.25 for a put option.

5.10.1 HJM driven by a Brownian motion

Table 5.2: Pricing parameters

Symbol	Meaning
$T = 1$	Option maturity
$U = 2$	Bond maturity
$\sigma = 0.015, a = 0.5$	Model parameters
$K = [0.9, 091, \cdots, 1]$	Strike prices
$f(0,t) = 0.05 \ \ \forall t$	Initial term structure

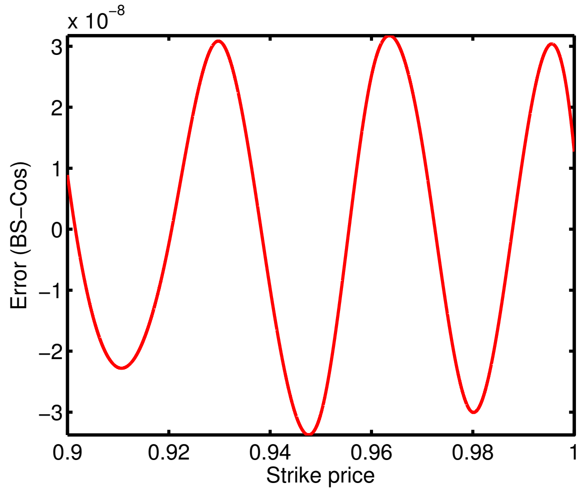
Source (Eberlein and Raible, 1999, p. 50)

Table 5.3: Option values for a HJM driven by a Brownian motion

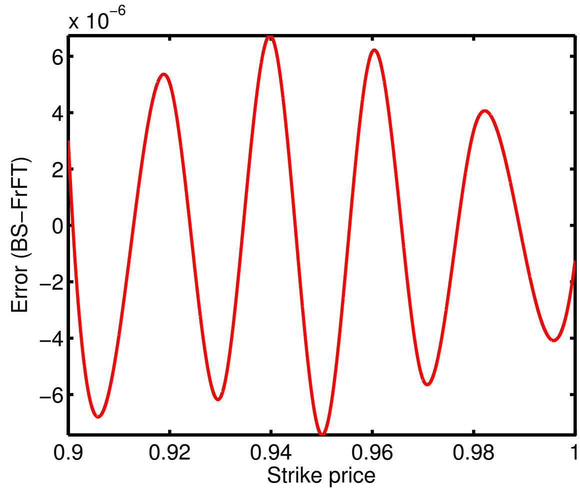
Pricing result for a HJM models driven by a Brownian motion						
	COS method		FFT method		FrFT method	
K	Call option	Put option	Call Option	Put option	Call option	Put option
0.9	0.0487	2.43×10^{-12}	0.0487	2.71×10^{-10}	4.8731×10^{-2}	2.4350×10^{-12}
0.91	0.0392	1.90×10^{-9}	0.0392	2.39×10^{-9}	3.9219×10^{-2}	1.9073×10^{-9}
0.92	0.0297	2.89×10^{-7}	0.0297	4.09×10^{-7}	2.9707×10^{-2}	3.8897×10^{-7}
0.93	0.0202	2.25×10^{-5}	0.0202	2.22×10^{-5}	2.0217×10^{-2}	2.2507×10^{-5}
0.94	0.0111	4.14×10^{-4}	0.0111	4.11×10^{-4}	1.1095×10^{-2}	4.1355×10^{-4}
0.95	0.0040	0.0028	0.0040	0.0028	4.0024×10^{-3}	2.8330×10^{-3}
0.96	7.41×10^{-4}	0.0091	7.39×10^{-4}	0.0091	7.4116×10^{-4}	9.0840×10^{-3}
0.97	5.83×10^{-5}	0.0179	5.63×10^{-5}	0.0179	5.8267×10^{-5}	1.7913×10^{-2}
0.98	1.75×10^{-6}	0.0274	1.68×10^{-6}	0.0274	1.7515×10^{-6}	2.7369×10^{-2}
0.99	1.92×10^{-8}	0.0369	2.36×10^{-8}	0.0369	1.9253×10^{-8}	3.6880×10^{-2}
1	7.64×10^{-11}	0.0464	-4.18×10^{-10}	0.0464	7.6369×10^{-11}	4.6392×10^{-2}

Stellenbosch University https://scholar.sun.ac.za

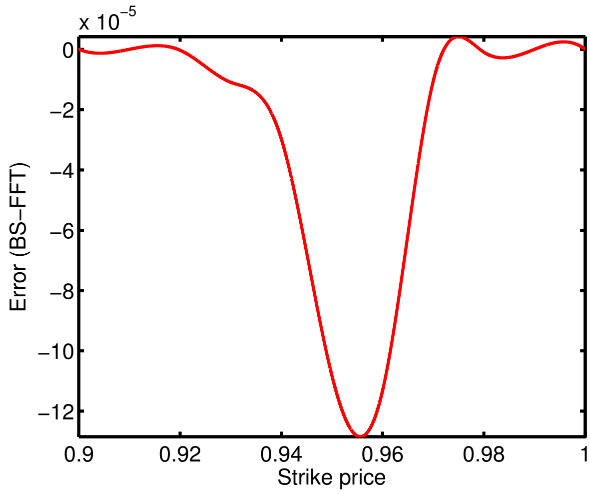
The plots in Figure 5.3 show errors in the methods compared to the analytical formulae in Equation 4.9.7.



(a) Cosine errors, a call option in Gaussian HJM



(b) FrFT errors, a call option in Gaussian HJM



(c) FFT errors, a call option in Gaussian HJM

Figure 5.3: Errors in Black–Scholes type formulae compared to various methods.

Stellenbosch University https://scholar.sun.ac.za

Table 5.4: Speed and convergence comparison between pricing methods. Reference value is obtained from Black–Scholes type formulae.

Method comparizon results for HJM model driven by a Brownian motion						
	COS method		FFT method		FrFT method	
N	Time (s)	RMSE (ϵ)	Time (s)	RMSE (ϵ)	Time (s)	RMSE (ϵ)
2^6	0.021265	0.0016579	0.025380	1.1579	0.017793	0.010411
2^7	0.020218	3.5537×10^{-4}	0.025877	0.012537	0.017681	0.0050766
2^8	0.027500	2.2919×10^{-5}	0.030358	0.0026377	0.019144	0.0012684
2^9	0.030232	2.0474×10^{-8}	0.029281	0.0014432	0.018945	2.2478×10^{-4}
2^{10}	0.033602	3.7292×10^{-16}	0.027080	6.3006×10^{-4}	0.019067	7.3121×10^{-6}
2^{11}	0.056831	5.4960×10^{-16}	0.030337	1.4275×10^{-4}	0.020761	8.5397×10^{-8}
2^{12}	0.069397	5.4960×10^{-16}	0.031103	4.8306×10^{-5}	0.024034	2.6999×10^{-8}
2^{13}	0.13982	5.4960×10^{-16}	0.036597	1.3385×10^{-5}	0.028628	8.2221×10^{-9}
2^{14}	0.26505	5.4960×10^{-16}	0.048641	2.7484×10^{-5}	0.040899	2.2780×10^{-9}

5.10.2 HJM driven by generalised hyperbolic motion

The characteristic function of a random variable X_T has the same form as the one in Equation 5.3.2, where the driving process is the GH.

Table 5.5: Pricing parameters for a model driven by a GH

Symbol	**Meaning**
$T = 1$	Option maturity
$U = 2$	Bond maturity
$\sigma = 1.5, a = 0.5, \alpha = 40$	
$\lambda = 0.5, \beta = -8, \delta = 0.1, \mu = 0$	Model parameters
$K = [0.9, 091, \cdots, 1]$	Strike prices
$f(0,t) = 0.05 \;\; \forall t$	Initial term structure

Stellenbosch University https://scholar.sun.ac.za

Table 5.6: Option values for HJM driven by GH

Pricing result for a HJM models driven by a GH						
	COS method		FFT method		FrFT method	
K	Call option	Put option	Call option	Put option	Call option	Put option
0.9	0.0529659	0.0042350	0.0529669	0.0042373	0.0529659	0.0042363
0.91	0.0450734	0.0058548	0.0450744	0.0058402	0.0450734	0.0058393
0.92	0.0376865	0.0079801	0.0376871	0.0079385	0.0376865	0.0079379
0.93	0.0309035	0.0107094	0.0309040	0.0106350	0.0309035	0.0106345
0.94	0.0248146	0.0141328	0.0248154	0.0140308	0.0248146	0.0140300
0.95	0.0194890	0.0183195	0.0194904	0.0182133	0.0194890	0.0182118
0.96	0.0149631	0.0233060	0.0149652	0.0232414	0.0149631	0.0232392
0.97	0.0112329	0.0290880	0.0112348	0.0291291	0.0112329	0.0291270
0.98	0.0082531	0.0356205	0.0082535	0.0358329	0.0082531	0.0358325
0.99	0.0059444	0.0428241	0.0059458	0.0432519	0.0059444	0.0432505
1	0.0042063	0.0505983	0.0042072	0.0512225	0.0042063	0.0512217

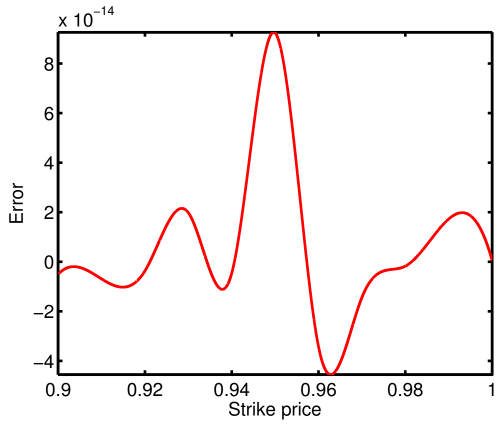


Figure 5.4: Cosine errors, a put option in GH HJM

Stellenbosch University https://scholar.sun.ac.za

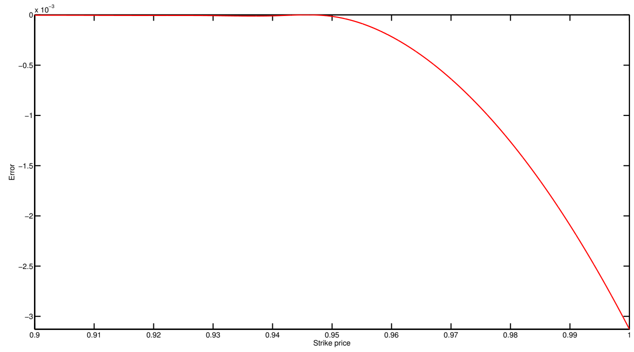


Figure 5.5: FFT errors, a put option in GH HJM

Table 5.7: Speed and convergence comparison between pricing methods. Reference values are obtain from numerical integration.

Method comparizon results for HJM model driven by a GH						
	COS method		FFT method		FrFT method	
N	Time (s)	RMSE (ϵ)	Time (s)	RMSE (ϵ)	Time (s)	RMSE (ϵ)
2^5	0.017290	0.0010542	0.027998	5164.5	0.020435	0.0085234
2^6	0.02043	4.7257×10^{-5}	0.028514	0.35000	0.020661	0.0026055
2^7	0.025649	1.3772×10^{-7}	0.028966	0.0087749	0.021491	2.9958×10^{-4}
2^8	0.036815	2.5651×10^{-12}	0.030390	0.0044304	0.022416	2.3686×10^{-4}
2^9	0.057766	0	0.032821	0.0013786	0.024632	2.4343×10^{-4}
2^{10}	0.10069	0	0.037260	4.4507×10^{-4}	0.029232	2.4336×10^{-4}
2^{11}	0.18256	0	0.046154	2.7972×10^{-4}	0.038336	2.4333×10^{-4}
2^{12}	0.34663	0	0.063296	2.4851×10^{-4}	0.055305	2.4333×10^{-4}
2^{13}	0.70226	0	0.095846	2.4463×10^{-4}	0.088365	2.4332×10^{-4}

5.11 Analysis of results

In this section, we discuss briefly the methods introduced in this chapter. We limit our discussion to FFT, FrFT and the COS method. All methods have a number of significant advantages, including the ability to value option under any model where the characteristic function is known. In addition their speed is motivational.

All methods can be implemented by setting the discrete summation grids to be $N = 2^n$, to speed up calculations. The approximation of FFT and FrFT depends on the discrete summation grid with Δu: the smaller the grid the better. However, in FFT, only two of the three parameters Δu,

Stellenbosch University https://scholar.sun.ac.za

Δx and N can be arbitrarily chosen; the third must satisfy $\Delta x = \frac{2\pi}{\Delta u N}$, the strike grid. Usually Δu and N are chosen to make the integration accurately enough, hence making the strike grid inversely proportional to the upper integration bound. This means the finer the grid Δu, the coarser the strike grid Δx, creating a trade-off between accuracy and the number of strikes. If the strike grid is wider, FFT generates strikes that are out of the region of interest.

A alternative improvement of FFT is the fractional fast Fourier transform (FrFT). It frees the integration grid choice from the strike grid, which makes the method more efficient. N-point FrFT procedures require a similar number of operations as a $4N$-point FFT procedure.

The COS method on the other hand requires $O(n)$ multiplications, which is computationally faster then the $O(n \log n)$ of the FFT. As a result, it can be viewed as an significant improvement of the FFT method. The only error in this method occurs when truncating the infinite range to the finite range. The COS prices are sensitive to the choice of the interval $[a, b]$. Hence, the COS method requires a proper estimate of $[a, b]$.

Implementing Fourier-based numerical methods requires choosing of a dampening factor R, which critically affects the pricing accuracy. We have examined and found out that for most stable values for $R \in (-\infty, 0)$ are the ones near 0 for a put option. This is true for Fourier methods. However, FFT and FrFT requires a bigger choice of R. Raible (2000) suggested that the optimal choice for R in the FFT method is $R = 25$.

Pricing methods such as the FFT, FrFT and COS methods are introduced for easy valuation of options on zero-coupon bonds. All these methods are implemented in MATLAB. Because the characteristic function involves the integral, we found that the six point Gauss–Legendre provides sufficiently accurate results.

When comparing numerical results of the comparison between the FFT, FrFT and COS methods both methods, produces the same results which implies that the models are well developed. The COS method converges fast, then FrFT and FFT. FrFT seems to be faster than the COS method in most cases. All three methods show that at-the-money options error forms a global minimum.

Stellenbosch University https://scholar.sun.ac.za

Chapter 6

Calibration

In the last two chapters, we discussed the Lévy HJM models and several numerical valuation methods. In chapter 5, we have implemented different methods for numerical evaluation assuming that model parameters are given. In general this pricing models depends on variables that are not directly observable. To make use of a model or to prove the validation of the model, the preliminary step is to determine unobservable variables. This is normally done either by backing them out, i.e. calibrating them, from liquid market data of derivatives or by estimating them.

The goal of this chapter is to present model calibration. This is a process of deriving unobserved model parameters by matching the model prices with the quoted market prices. It involves continuous adjustments of the model parameters until the model prices are compatible with the quoted market prices.

Strictly speaking, to be able to use any financial model, one has to specify the model parameters by means of calibration. The main aim of calibration is to obtain (back-out) unobserved model parameters that describe the current state of the market. Calibration of the term structure model to represent the current state is one of a model's most important features. Model calibration is a crucial step in option pricing theory.

In Chapter 4, we assumed that the model parameters were known when we valued options on zero-coupon bonds. However, in order to make the model relevant to real-world application, we must carry out model calibration to obtain optimal parameters that best describes the current market state.

We are confronted with liquid quoted prices for caps/floors and swaptions from the South African fixed income market for a specific day and we would like to calibrate the Lévy HJM model as discussed in the previous chapters. The calibration is performed on European caps/floors and swaptions (payer) with maturities from 3 months up to 10 years.

We shall begin with the derivation of interest rate pricing formulae in Section 6.1. In section

Stellenbosch University https://scholar.sun.ac.za

6.2, we present the data acquired from local South African bank. Section 6.3 presents the calibration methods and results.

6.1 Pricing formulae for interest rate derivatives

6.1.1 Interest rate caps/floors

Recall from Equation 2.3.3, that $\tilde{\kappa} = \frac{1}{1+\delta_i \kappa}$ denotes a modified strike price. Adopting the usual notation, the pay-off function for a caplet is given by $\Phi(P(T_{i-1}, T_i)) = \frac{1}{\tilde{\kappa}} \max\{\tilde{\kappa} - P(T_{i-1}, T_i), 0\}$, and $\chi_{X_{T_{i-1}}}$ is the risk-neutral characteristic function in Equation 5.3.2 of a bond price driving process X_t (non-homogeneous Lévy process) under the T_{i-1}-forward measure. Based on Theorem 5.3.4, we have the following:

$$\begin{aligned} \mathbb{V}_{\text{caplet}}(t, T_{i-1}, T_i) &= \frac{1}{\tilde{\kappa}} P(t, T_{i-1}) \mathbb{E}_{\mathbb{Q}^{T_{i-1}}} \left[(\tilde{\kappa} - P(T_{i-1}, T_i))^+ \right] \\ &= \frac{P(0, T_{i-1})}{2\pi\tilde{\kappa}} e^{-R\xi_i} \int_{\mathbb{R}} e^{-\mathrm{i}u\xi_i} \hat{\Phi}(\mathrm{i}R - u) \chi_{X_{T_{i-1}}}(u - iR)\, du, \quad \forall R \in (-\infty, 0), \end{aligned} \tag{6.1.1}$$

and

$$\begin{aligned} \mathbb{V}_{\text{floorlet}}(t, T_{i-1}, T_i) &= \frac{1}{\tilde{\kappa}} P(t, T_{i-1}) \mathbb{E}_{\mathbb{Q}^{T_{i-1}}} \left[(P(T_{i-1}, T_i) - \tilde{\kappa})^+ \right] \\ &= \frac{P(0, T_{i-1})}{2\pi\tilde{\kappa}} e^{-R\xi_i} \int_{\mathbb{R}} e^{-\mathrm{i}u\xi_i} \hat{\Phi}(\mathrm{i}R - u) \chi_{X_{T_{i-1}}}(u - iR)\, du, \quad \forall R \in (1, +\infty), \end{aligned} \tag{6.1.2}$$

where $\xi_i = -\log D(T_{i-1}, T_i)$, $D(T_{i-1}, T_i) = \frac{P(0,T_i)}{P(0,T_{i-1})} \exp\left(\int_0^{T_{i-1}} \psi_i(S(t, T_{i-1})) - \psi_i(S(t, T_i))\, dt \right)$ and $\hat{\Phi}$ denotes the Fourier transform of a pay-off function.

Based on the definition for a cap and a floor (see Section 2.3), it follows that

$$\mathbb{V}_{\text{cap}}(t, T_j, \kappa) = \sum_{i=1}^{j} \mathbb{V}_{\text{caplet}}(t, T_{i-1}, T_i) \text{ and } \mathbb{V}_{\text{floor}}(t, T_j) = \sum_{i=1}^{j} \mathbb{V}_{\text{floorlet}}(t, T_{i-1}, T_i).$$

6.1.2 Swaptions

Consider the volatility structure with restriction given in Condition 4.5.1 and Condition 4.3.4. Notice that one can write

$$S(t, T_i) - S(t, T) = \sigma_1(t) \int_T^{T_i} \sigma_2(u)\, du, \text{ for } i = 1, 2, \cdots, n \tag{6.1.3}$$

and it follows that

$$S(t, T_n) - S(t, T) = \int_T^{T_n} \sigma(t, u)\, du = \sigma_1(t) \int_T^{T_n} \sigma_2(u)\, du. \tag{6.1.4}$$

Stellenbosch University https://scholar.sun.ac.za

Divide (6.1.3) by (6.1.4),

$$S(t,T_i) - S(t,T) = \frac{\int_T^{T_i} \sigma_2(u)\, du}{\int_T^{T_n} \sigma_2(u)\, du} \left(S(t,T_n) - S(t,T)\right). \tag{6.1.5}$$

Recall Equation 4.4.3 and denotes a random variable

$$\begin{aligned} X_{T_i} = \int_0^T S(s,T_i) - S(s,T)\, dL_s &= \int_0^T \frac{\int_T^{T_i} \sigma_2(u)\, du}{\int_T^{T_n} \sigma_2(u)\, du} \left(S(t,T_n) - S(t,T)\right)\, dL_s \\ &= \underbrace{\frac{\int_T^{T_i} \sigma_2(u)\, du}{\int_T^{T_n} \sigma_2(u)\, du}}_{R(T,T_i)} \overbrace{\int_0^T S(t,T_n) - S(t,T)\, dL_s}^{X_{T_n}}. \end{aligned}$$

Note: The volatility form restriction in Condition 4.5.1 is necessary for the decomposition of a random variable X_{T_i} such that

$$X_{T_i} = R(T,T_i) X_{T_n}.$$

Recalling Equation 4.4.3, a zero-coupon bond can be expressed in the form of

$$P(t,T_i) = D(t,T_i) \exp\left(R(t,T_i) X_{T_n}\right).$$

We assume that the distribution of X_{T_n} possesses a density function under the forward martingale measure $\mathbb{Q}^T$. This enables us to calculate the option price via transform methods, (see Raible, 2000, Proposition 4.4). Another approach is to verify Conditions 5.3.3.

Rewrite Equation 2.3.6 as

$$\mathbb{V}_{\text{PS}}(0) = P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[\left(1 - \sum_{i=1}^{n} c_i D(T,T_i) e^{R(T,T_i)X_{T_n}}\right)^+ \Bigg| \mathcal{F}_t\right]. \tag{6.1.6}$$

Define a function $f(s,t,x) = D(s,t)e^{R(s,t)x}$. Realise that due to the volatility factorisation in Condition 4.5.1, $f(T,T_i,x)$ is an increasing function of x for all $i = 1, 2 \cdots n$, since $D(T,T_i), R(T,T_i) > 0$, therefore the Jamshidian (1989) decomposition (see Appendix A.4) is valid:

$$\mathbb{V}_{\text{PS}}(t) = P(0,T) \sum_{i=1}^{n} c_i \mathbb{E}_{\mathbb{Q}^T}\left[\left(K_i - f(T,T_i,X_{T_n})\right)^+ \big| \mathcal{F}_t\right], \tag{6.1.7}$$

where $K_i = f(T,T_i,K)$ for unique K that solves the equation

$$\sum_{i=1}^{n} c_i f(T,T_i,K) = 1. \tag{6.1.8}$$

According to Jamshidian (1989), we have the following:

Stellenbosch University https://scholar.sun.ac.za

Theorem 6.1.1 *The price of a payer swaption is given by*

$$\mathbb{V}_{PS}(0) = P(0,T)\sum_{i=1}^{n} c_i \mathbb{V}^i_{put}(0,T,T_i,K_i), \tag{6.1.9}$$

where $\mathbb{V}^i_{put}(0,T,T_i,K_i)$ is a put option value with maturity T, strike price K_i on a zero-coupon bond maturing at time T_i given by

$$\mathbb{V}_{put}(0,T,U,K) = \frac{e^{-R\xi}}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}u\xi}\hat{\Phi}(\mathrm{i}R-u)\chi_{X_{T_n}}(u-iR)\,du, \quad \forall R \in (-\infty, 0), \tag{6.1.10}$$

$\xi = -\log D(T,U)$.

The result follows from Theorem 5.3.4 (also see the appendix A.5).

Eberlein and Kluge (2006) used a different approach. We review their approach briefly.

$$\mathbb{V}_{\mathrm{PS}}(t) = P(0,T)\mathbb{E}_{\mathbb{Q}^T}\left[\left(1-\sum_{i=1}^{n}\tilde{D}(T,T_i)e^{R(T,T_i)X_{T_n}}\right)^{+}\middle|\mathcal{F}_t\right], \tag{6.1.11}$$

where $\tilde{D}(T,T_i) = c_i D(T,T_i)$ and $R(T,T_i) = \frac{\int_T^{T_i}\sigma_2(u)\,du}{\int_T^{T_n}\sigma_2(u)\,du}$.

Denote

$$g(s,t,x) = 1-\sum_{i=1}^{n}\tilde{D}(s,t_i)e^{R(s,t_i)x}.$$

The assumption made for the volatility structure guarantees that the map $x \mapsto g(T,T_i,x)$ is an increasing function of x for all $i = 1,2,\cdots,n$. There exist a unique K such that $g(s,t,K) = 0$.

The following theorem by (Eberlein and Kluge, 2006, Theorem 16) finds the value of a payer swaption:

Theorem 6.1.2 *Suppose the distribution of X_{T_n} with respect to $\mathbb{Q}^T$ possesses a density function. Choose an $R<0$ such that $\chi_{X_{T_n}}(\mathrm{i}R) < \infty$ and let K be the unique zero of the strictly increasing and continuous function*

$$g(s,t,x) = 1-\sum_{i=1}^{n}\tilde{D}(s,t_i)e^{R(s,t_i)x}.$$

Then,

$$\mathbb{V}_{PS}(t) = \frac{P(0,T)}{2\pi}\int_{\mathbb{R}} L[v](R+\mathrm{i}u)\chi_{X_{T_n}}(u-\mathrm{i}R)\,du, \quad \forall R \in (-\infty,0), \tag{6.1.12}$$

where

$$L[v](u) = e^{uK}\left(\frac{1}{u}-\sum_{i=1}^{n}\left[\tilde{D}(T,T_i)e^{R(T,T_i)K}\frac{1}{R(T,T_i)+u}\right]\right)$$

denotes the bilateral Laplace transform of a function $v:\mathbb{R}\mapsto\mathbb{R}$ defined by $v(x) = (-g(s,t,-x))^{+}$.

Stellenbosch University https://scholar.sun.ac.za

6.2 Data

The dataset consists of ATM caps/floors and swaptions volatility market data from 8 September 2013. More specifically, the data consists of the South African 3 month JIBOR rate, with maturities ranging from three-months to 30 years.[1]

Table 6.1: JIBOR swaption implied volatilities on 8 September 2013

Swaption ATM Vols											
Exp/Tenor	1yr	2yr	3yr	4yr	5yr	6yr	7yr	8yr	9yr	10yr	50yr
0	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19
1	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19
3	0.205	0.215	0.205	0.1975	0.19	0.1875	0.185	0.183	0.182	0.18	0.18
6	0.205	0.22	0.22	0.22	0.22	0.2175	0.215	0.213	0.212	0.21	0.21
9	0.21	0.22	0.22	0.22	0.22	0.2175	0.215	0.213	0.212	0.21	0.21
12	0.23	0.24	0.23	0.225	0.22	0.2175	0.215	0.213	0.212	0.21	0.21
24	0.24	0.235	0.23	0.2275	0.225	0.22125	0.2175	0.215	0.2125	0.21	0.21
36	0.24	0.23	0.23	0.2275	0.225	0.22125	0.2175	0.215	0.2125	0.21	0.21
48	0.235	0.225	0.225	0.22375	0.2225	0.21875	0.215	0.2125	0.210	0.208	0.208
60	0.23	0.22	0.22	0.22	0.22	0.21625	0.2125	0.21	0.2075	0.205	0.205
72	0.228	0.22	0.218	0.2175	0.217	0.214	0.211	0.209	0.207	0.205	0.205
84	0.226	0.22	0.216	0.215	0.214	0.21175	0.2095	0.208	0.2065	0.205	0.205
96	0.224	0.22	0.214	0.2125	0.211	0.2095	0.208	0.207	0.206	0.205	0.205
108	0.222	0.22	0.212	0.21	0.208	0.20725	0.2065	0.206	0.2055	0.205	0.205
120	0.22	0.22	0.21	0.2075	0.205	0.205	0.205	0.205	0.205	0.205	0.205

[1]These data were provided by Phetha Ndlangamandla from the South African Rand Merchant Bank.

Stellenbosch University https://scholar.sun.ac.za

Table 6.2: Money-market instruments

Instrument object	Indicator	Fixed rate (%)
Forward rate Agreement (FRA)		
FRA.ZARJIBAR3M_FRA1X4	TRUE	5.91%
FRA.ZARJIBAR3M_FRA2X5	TRUE	6.13%
FRA.ZARJIBAR3M_FRA3X6	TRUE	6.19%
FRA.ZARJIBAR3M_FRA4X7	TRUE	6.40%
FRA.ZARJIBAR3M_FRA5X8	TRUE	6.46%
FRA.ZARJIBAR3M_FRA6X9	TRUE	6.69%
FRA.ZARJIBAR3M_FRA7X10	TRUE	6.74%
FRA.ZARJIBAR3M_FRA8X11	TRUE	6.89%
FRA.ZARJIBAR3M_FRA9X12	TRUE	6.95%
FRA.ZARJIBAR3M_FRA12X15	TRUE	7.26%
FRA.ZARJIBAR3M_FRA15X18	TRUE	7.53%
FRA.ZARJIBAR3M_FRA18X21	TRUE	7.73%
Fixed Rate: References an Excel error.	FALSE	Err:512
Swap rates		
INTERESTRATESWAP.ZARJIBAR3M_SWAP1Y	FALSE	6.36%
INTERESTRATESWAP.ZARJIBAR3M_SWAP2Y	TRUE	6.97%
INTERESTRATESWAP.ZARJIBAR3M_SWAP3Y	TRUE	7.36%
INTERESTRATESWAP.ZARJIBAR3M_SWAP4Y	TRUE	7.63%
INTERESTRATESWAP.ZARJIBAR3M_SWAP5Y	TRUE	7.84%
INTERESTRATESWAP.ZARJIBAR3M_SWAP6Y	TRUE	8.02%
INTERESTRATESWAP.ZARJIBAR3M_SWAP7Y	TRUE	8.15%
INTERESTRATESWAP.ZARJIBAR3M_SWAP8Y	TRUE	8.26%
INTERESTRATESWAP.ZARJIBAR3M_SWAP9Y	TRUE	8.35%
INTERESTRATESWAP.ZARJIBAR3M_SWAP10Y	TRUE	8.43%
INTERESTRATESWAP.ZARJIBAR3M_SWAP12Y	TRUE	8.54%
INTERESTRATESWAP.ZARJIBAR3M_SWAP15Y	TRUE	8.64%
INTERESTRATESWAP.ZARJIBAR3M_SWAP20Y	TRUE	8.72%
INTERESTRATESWAP.ZARJIBAR3M_SWAP25Y	TRUE	8.69%
INTERESTRATESWAP.ZARJIBAR3M_SWAP30Y	TRUE	8.66%
Fixed Rate: References an Excel error.	FALSE	Err:512

Stellenbosch University https://scholar.sun.ac.za

Table 6.3: Euro caplet implied volatilities on 8 September 2013

Caps/floor ATM Vols									
Tenor	Current	Tenor	Current	Tenor	Current	Tenor	Current	Tenor	Current
0	0.165	*24*	0.312	*48*	0.236	*72*	0.250	*96*	0.245
3	0.165	*27*	0.312	*51*	0.236	*75*	0.250	*99*	0.245
6	0.204	*30*	0.312	*54*	0.236	*78*	0.250	*102*	0.245
9	0.228	*33*	0.312	*57*	0.236	*81*	0.250	*105*	0.245
12	0.25	*36*	0.254	*60*	0.250	*84*	0.245	*108*	0.245
15	0.26	*39*	0.254	*63*	0.250	*87*	0.245	*111*	0.245
18	0.275	*42*	0.254	*66*	0.250	*90*	0.245	*114*	0.245
21	0.304	*45*	0.254	*69*	0.250	*93*	0.245	*117*	0.245

Table 6.3, 6.1 and 6.2 presents the quoted caps and swaptions volatilities, and money market instruments respectively. Note that prices are not measured in nominal values, but rather in implied Black volatilities, as is the market convention. To obtain market prices, we need to calculate them from Black volatilities using Black's formula:

$$\mathbb{V}^{\mathrm{M}}_{\text{caplet}}(0, T_i, T_{i+1}, \kappa) = P(0, T_i)\tau\left(f(0, T_i, T_{i+1})N(d_1) - \kappa N(d_2)\right) \tag{6.2.1}$$

$$d_1 = \frac{\log \frac{f(0,T_i,T_{i+1})}{\kappa} + \frac{1}{2}\sigma^2_{\text{caplet}} T_i}{\sigma_{\text{cap}}\sqrt{T_i}}, \quad d_2 = d_1 - \sigma_{\text{caplet}}\sqrt{T_i},$$

d_1 and d_2 depend on T_i and T_{i+1} where

$$f(0, T_i, T_{i+1}) = -\frac{1}{\tau}\log\left[\frac{P(0, T_{i+1})}{P(0, T_i)}\right] \tag{6.2.2}$$

is the forward LIBOR rate, and τ is the tenor and σ_{caplet} is the market caplet volatility obtained from stripping caps/floors volatilities. Cap prices are obtained by

$$\mathbb{V}^{\mathrm{M}}_{\text{cap}}(0, T_i, \kappa) = \sum_{k=1}^{i} \mathbb{V}^{\mathrm{M}}_{\text{caplet}}(0, T_k, T_{k+1}, \kappa).$$

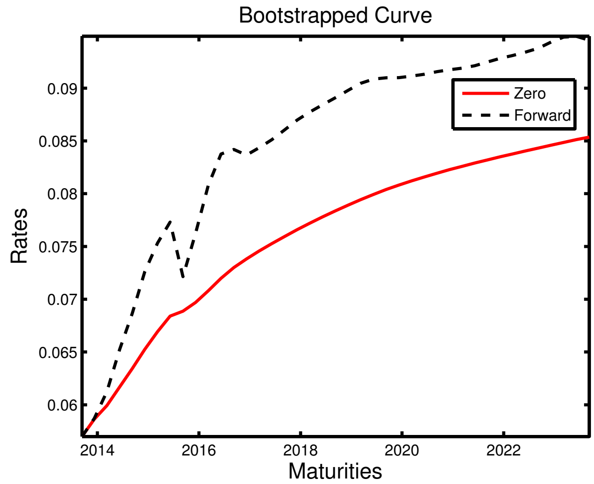
To compute payer swaption prices, we used the following Black formula:

$$\mathbb{V}_{\text{PS}} = (S_{0n}N(d_1) - \kappa N(d_2))\sum_{i=1}^{n}\delta_i P(0, T_i)$$

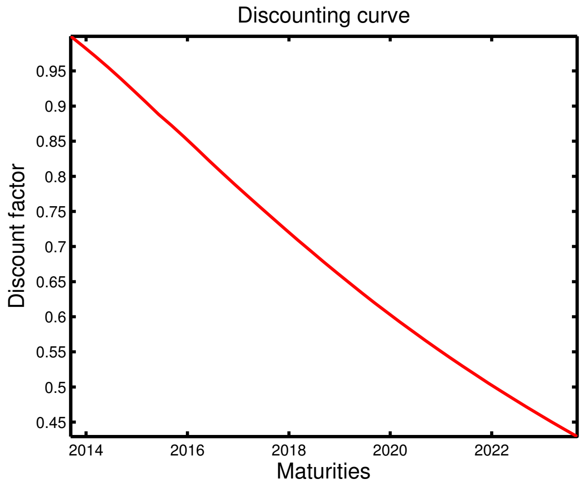
$$d_1 = \frac{\log \frac{S_{0n}}{\kappa} + \frac{1}{2}\sigma^2_{\text{swaption}} T_0}{\sigma_{\text{swaption}}\sqrt{T_0}}, \quad d_2 = d_1 - \sigma_{\text{swaption}}\sqrt{T_0}, \tag{6.2.3}$$

where S_{0n} is a forward start swap rate on a forward swap starting at time $T = 0$ and maturing at time T_n.

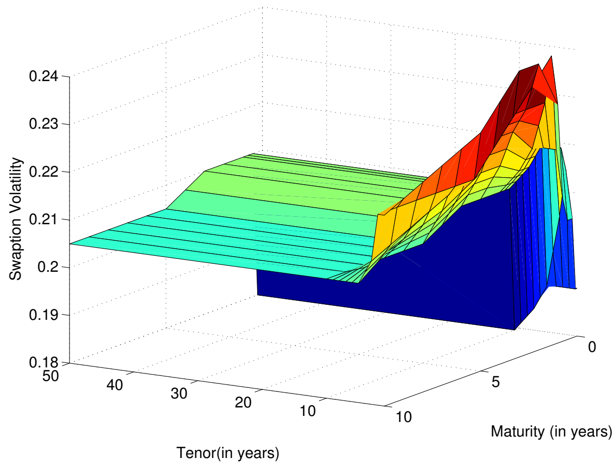
Stellenbosch University https://scholar.sun.ac.za



(a) Zero-curve stripped from market JIBOR rates on 8 September 2013.



(b) Term structure of discount factors stripped from market JIBOR rates on 8 September 2013.



(c) SA rand swaption volatility surface

Figure 6.1: Interest rate quoted data on 8 September 2013

Figures 6.1a, 6.1b and 6.1c show the yield curve, the discount bond and swaption volatility surface. Data for Figure 6.1a and 6.1b have been obtained by bootstrapping the quoted par interest rates (FRA and IRS) data in Table 6.2. There is a hump peaking at roughly three to four year maturity while the volatility stays constant for greater maturities, such as five to ten years.

Stellenbosch University https://scholar.sun.ac.za

6.3 Calibration methods

The calibration procedure minimise the differences between model prices and market quoted prices. Finally, the objective is to obtain the best parameters such that these differences is minimised. The calibration problem is summarised below:

Problem 6.3.1 (Calibration) *Let $\Theta \subset \mathbb{R}$ be a set of unobserved model parameters and $H : \Theta \mapsto \mathbb{R}$ be a pricing function. The calibration problem seeks the global minimum m,*

$$m = \inf_{\Theta} f[\Theta], \tag{6.3.1}$$

where the objective function $f[\Theta]$ takes the form

(a)

$$f[\Theta] = \frac{1}{\sqrt{n}} \sqrt{\sum_{i=0}^{n} |MP_i - H_i(\Theta)|^2},$$

(b)

$$f[\Theta] = \frac{|H_i(\Theta) - MP_i|^2}{MP_i},$$

$n \in \mathbb{Z}$ denotes the number quoted option prices under consideration, MP_i is the market price.

The calibration problem above finds a vector Θ of model parameters in a such a way that the model price best fits the available quoted market prices. Hence, the problem can be interpreted as an optimisation problem, see (Kienitz and Wetterau, 2012, Chapter 9).

This study considers two methods of optimisation for the calibration problem above, namely simulated Annealing and secant-Levenberg–Marquardt method. The former is a stochastic optimisation method while the latter is based on non-linear optimisation with derivatives.

6.3.1 Simulated Annealing

This method is a global optimisation method that has attracted significant attention because it is suitable for optimisation problems on a large scale, especially ones where a desired global extreme is hidden among many, poorer, local extrema. It was introduced in the field of thermodynamics to control the dynamics of freezing and crystallising liquids (see Kirkpatrick *et al.* (1983)).

6.3.2 Secant-Levenberg–Marquardt

This method is an improved version of the Levenberg–Marquardt method for identifying a solution for a non-linear least squares curve fitting problem. It is an iterative technique that

Stellenbosch University https://scholar.sun.ac.za

finds the minimum of a function that is expressed as the sum of squares of non-linear functions. It is a combination of the gradient secant method and the Gauss–Newton method.

For more detailed theory on these methods, (see Kienitz and Wetterau, 2012, Chapter 9).

6.4 Calibration results

This section presents the calibration results of the HJM model driven by a Brownian motion and by a generalised hyperbolic process to market price of caps.

Before we start with the calibration we shall investigate the risk-neutral density to observe which parameters are necessary and which ones are redundant in the models.

6.4.1 Risk-neutral densities

This section illustrates changes in the risk neutral density as model parameters are varied. By changing the model parameters we expect change in the distribution since the distribution is responsible for the calculation of option values. The main aim is to observe which parameters contribute most to the option price.

One of the main Lévy HJM model's components is the volatility structure. From the theory chapter, we have opted to make use of Vasiček forward volatility given by

$$\sigma(t,T) = \hat{\sigma} e^{-a(T-t)}, \quad \hat{\sigma} > 0, \ a \neq 0.$$

Therefore the bond volatility structure is given by

$$S(t,T) = \int_t^T \sigma(t,s)\, ds = \frac{\hat{\sigma}}{a}\left(1 - e^{-a(T-t)}\right), \quad \hat{\sigma} > 0, \ a \neq 0.$$

Recall from Equation 5.3.2 that the characteristic function of a random variable X_T under a T-forward martingale measure is given by

$$\chi_{X_T}(u) = \exp\left(\int_0^T f(s) ds\right),$$

where the integrand is given by

$$f(s) = \psi_s\left(\mathrm{i}uS(s,U) + (1-\mathrm{i}u)S(s,T)\right) - \psi_s\left(S(s,T)\right).$$

Consider a case where a driving process is the generalised hyperbolic process, i.e.,

$$\psi(u) = \log\left[e^{\mu u}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+u)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda(\delta\sqrt{\alpha^2-(\beta+u)^2})}{K_\lambda(\delta\sqrt{\alpha^2-\beta^2})}\right].$$

Stellenbosch University https://scholar.sun.ac.za

Basically, we have to evaluate $\psi\left(\mathrm{i}uS(s,U)+(1-\mathrm{i}u)S(s,T)\right)$ and $\psi\left(S(s,T)\right)$. The latter becomes:

$$\psi\left(S(s,T)\right)=\log\left[e^{\mu S(s,T)}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+S(s,T))^2}\right)^{\frac{\lambda}{2}}\frac{K_\lambda(\delta\sqrt{\alpha^2-(\beta+S(s,T))^2}}{K_\lambda(\delta\sqrt{\alpha^2-\beta^2})}\right]. \tag{6.4.1}$$

By the change of variables, let $\lambda=\tilde{\lambda}$, $\alpha=\tilde{\alpha}\hat{\sigma}$, $\delta=\frac{\tilde{\delta}}{\hat{\sigma}}$, $\mu=\tilde{\mu}\hat{\sigma}$ and $\beta=\tilde{\beta}\hat{\sigma}$. It is shown that a generalised hyperbolic distribution is closed under affine transformation (see Eberlein, 2001) which means that Equation 6.4.1 becomes

$$\begin{aligned}\psi\left(S(s,T)\right)&=\log\left[e^{\mu S(s,T)}\left(\frac{\hat{\sigma}^2\tilde{\alpha}^2-\hat{\sigma}^2\tilde{\beta}^2}{\hat{\sigma}^2\tilde{\alpha}^2-(\hat{\sigma}\tilde{\beta}+S(s,T))^2}\right)^{\frac{\tilde{\lambda}}{2}}\frac{K_{\tilde{\lambda}}(\frac{\tilde{\delta}}{\hat{\sigma}}\sqrt{\hat{\sigma}^2\tilde{\alpha}^2-(\hat{\sigma}\tilde{\beta}+\sum(s,T))^2}}{K_{\tilde{\lambda}}(\frac{\tilde{\delta}}{\hat{\sigma}}\sqrt{\hat{\sigma}^2\tilde{\alpha}^2-\hat{\sigma}^2\tilde{\beta}^2})}\right]\\&=\tilde{\mu}\sum(s,T)+\frac{\tilde{\lambda}}{2}\log\left(\frac{\tilde{\alpha}^2-\tilde{\beta}^2}{\tilde{\alpha}^2-(\tilde{\beta}+\sum(s,T))^2}\right)+\log\left[\frac{K_{\tilde{\lambda}}(\tilde{\delta}\sqrt{\tilde{\alpha}^2-(\tilde{\beta}+\sum(s,T))^2}}{K_{\tilde{\lambda}}(\tilde{\delta}\sqrt{\tilde{\alpha}^2-\tilde{\beta}^2})}\right],\end{aligned}$$

where $\sum(s,T)=\frac{1}{a}\left(1-e^{-a(T-s)}\right)$. The same factorization holds for

$$\psi\left(\mathrm{i}uS(s,U)+(1-\mathrm{i}u)S(s,T)\right).$$

A priori, we have shown that a parameter $\hat{\sigma}$ can be a multiplicative constant in the driving process, therefore without loss of generality it can be chosen to be equal to 1. Furthermore, we found out that if parameter $0<\hat{\sigma}<1$ is not set to one, then the whole price process will merely depends on the given value of $\hat{\sigma}$, making other parameters unnecessary. Hence for the HJM model driven by a GH process we have six parameters $a,\lambda,\alpha,\beta,\delta,\mu$. We showed further that μ has no effect on the price process by looking at the risk-neutral densities, and hence it can be set to 0.

We have chosen an option maturing in one year ($T=1$) on a zero-coupon bond with a maturity of two years ($U=1.25$).

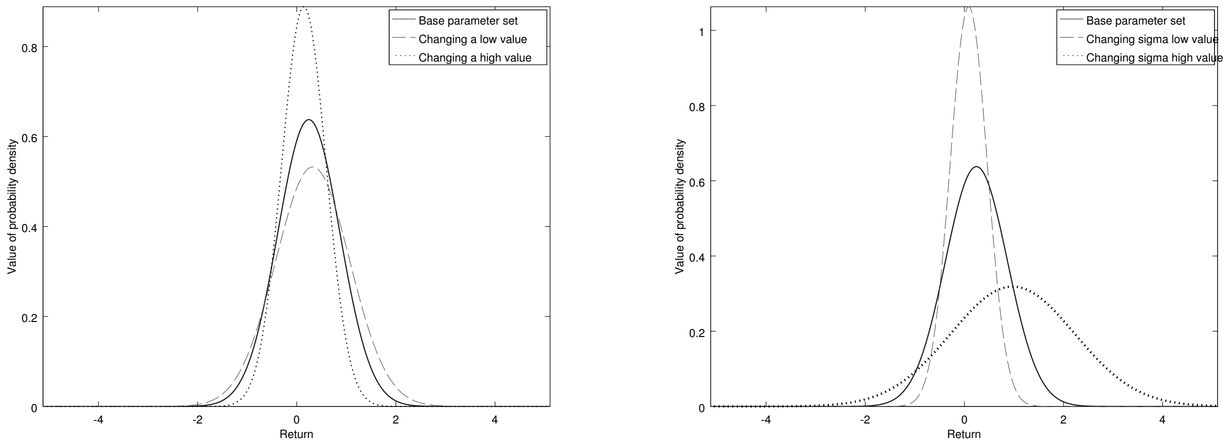
Using Fourier inversion theorem, one can derive the risk-neutral probability density from the given characteristic function of a random variable. Since the characteristic function is not known in closed form, we opted to use the six-points Gauss–Legendre integration and we obtained fairly satisfactory results.

6.4.2 Gaussian HJM

Gaussian HJM is referred to as a Lévy HJM driven by a Brownian motion under deterministic forward volatility.

Base parameter: $\sigma=1.5, a=0.5$.

Stellenbosch University https://scholar.sun.ac.za



(a) changing parameter a: low$=0.3$ and high$=0.9$ (b) changing parameter σ: low$=0.6$ and high$=2$

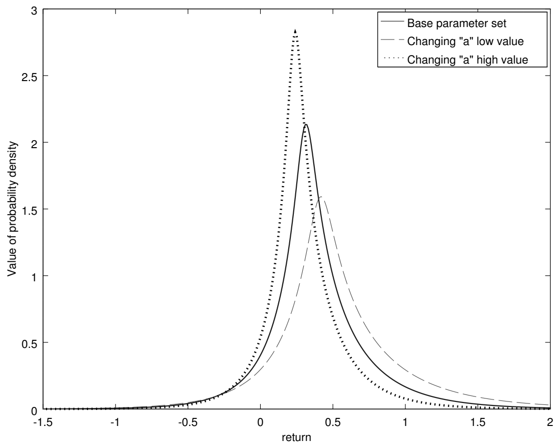
Figure 6.2: Risk-neutral density for a model driven by a Brownian motion

Stellenbosch University https://scholar.sun.ac.za

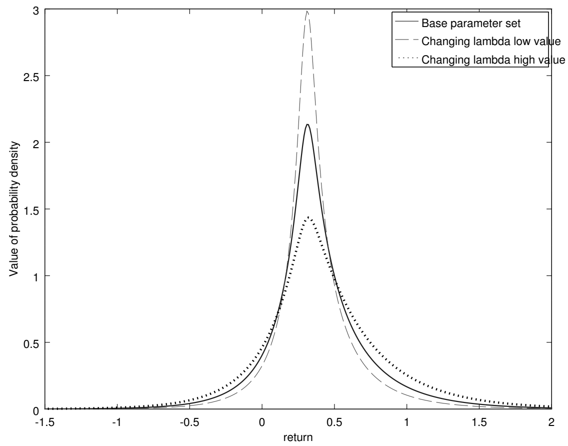
6.4.3 General HJM

We consider a HJM driven by a generalised hyperbolic motion.

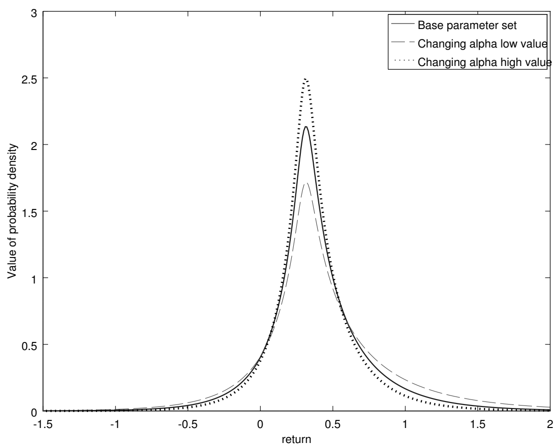
Base parameter set: $a = 0.5, \lambda = 0.5, \alpha = 2, \beta = 0, \delta = 0.1, \mu = 0.5$



(a) changing parameter a: low$= 0.2$ and high$= 0.8$



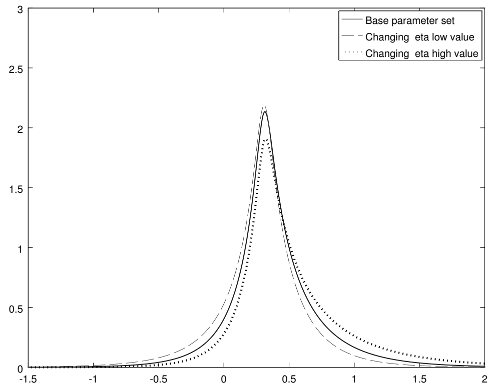
(b) changing parameter λ: low $= 0$ and high $= 0.9$



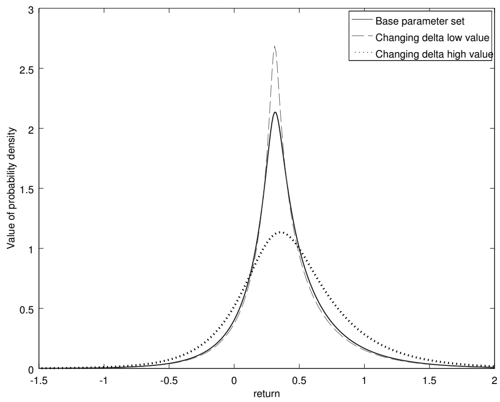
(c) Varrying parameter α: low $= 1.5$ and high $= 2.5$

Figure 6.3: Risk-neutral density for a model driven by GH; λ, α

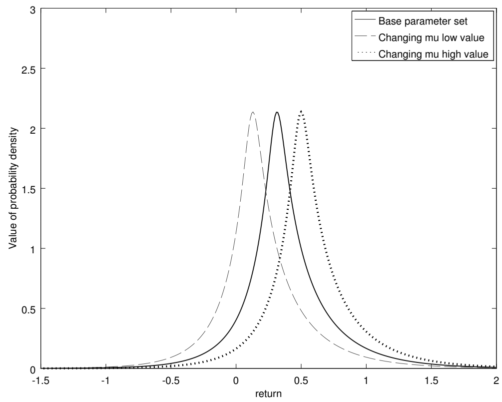
Stellenbosch University https://scholar.sun.ac.za



(a) changing parameter β: low = -0.5 and high $= 0.5$



(b) changing parameter δ: low = 0 and high $= 0.2$



(c) changing parameter μ: low = 0.2 and high $= 0.8$

Figure 6.4: Risk-neutral density for a model driven by GH; β, δ, μ

It is clear from Figure 6.4c that the parameter μ has no effect on option price, hence we exclude it out, i.e. we set $\mu = 0$.

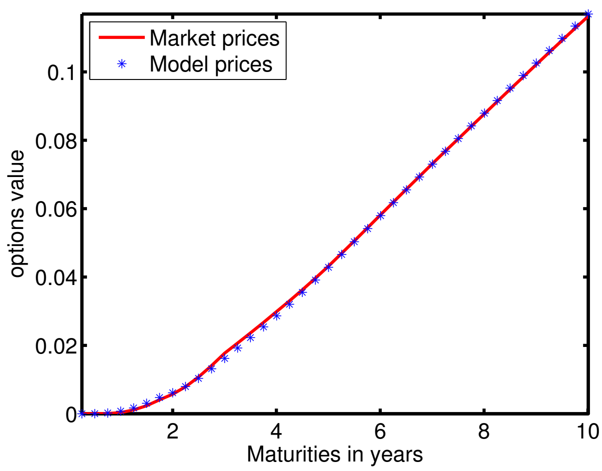
Stellenbosch University https://scholar.sun.ac.za

Table 6.4: Calibration results: Secant-Levenberg–Marquardt

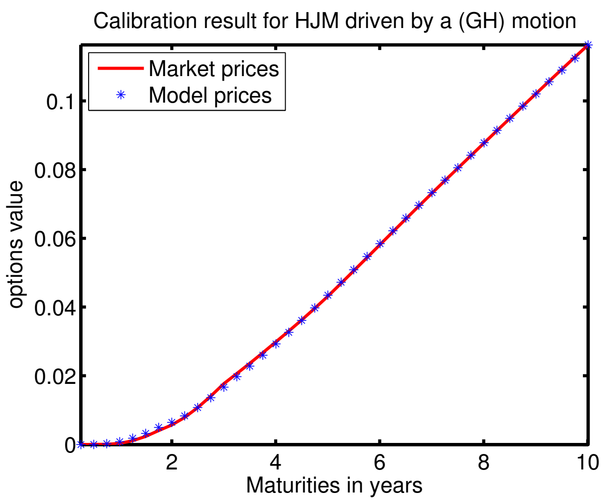
Secant-Levenberg–Marquardt		
Model driver:	**Brownian motion**	**Generalised hyperbolic**
Iniatial parameters	$\hat{\sigma} = 0.5, a = 0.9$	$a = 1.5, \lambda = 0.3$ $\alpha = 16, \beta = -4, \delta = 0.09$
Calibrated parameters	$\hat{\sigma} = 0.0241, a = 0.0396$	$a = 0.039139, \lambda = -6.857548$ $\alpha = 16.435150, \beta = -6.586567$ $\delta = 0.083660$
Time taken (s) Object function	35.68 $2.5081 \times e^{-5}$	2293.62 $4.6807 \times e^{-6}$

Table 6.5: Calibration results: Simulated Annealing

Simulated annealing		
Model driver:	**Brownian motion**	**Generalized hyperbolic**
Initial parameters	$\hat{\sigma} = 0.5, a = 0.9$	$a = 2, \lambda = 0.3$ $\alpha = 16, \beta = -4, \delta = 0.09$
Calibrated parameters	$\hat{\sigma} = 0.0241, a = 0.396$	$a = 0.040119, \lambda = -6.9290$ $\alpha = 16.3816, \beta = -5.1961$ $\delta = 0.09233$
Time taken (s) Object function	269.75 $5.0093 \times e^{-4}$	254320.416708 $3.89933207585559 \times e^{-5}$



(a) Calibration results for HJM driven by a Brownian motion



(b) Calibration results for HJM driven by a GH

Stellenbosch University https://scholar.sun.ac.za

6.5 Result discussion

In this chapter, we carried out calibration of Lévy HJM to the ATM caps prices using two methods of optimisation. Tables 6.4 and 6.5 display the calibration outcomes from these two methods. We obtained nearly the same parameters from both methods which proves the calibration stability. The simulated Annealing methods took longer than the secant-Levenberg–Marquardt. The calibration results for HJM driven by a Brownian motion and GH are shown in Figures 6.5a and 6.5b respectively. Based on the calibration results, both drivers provided a nearly perfect fit. Furthermore, the GH process provides a much better fit than a model driven by a Brownian motion.

Stellenbosch University https://scholar.sun.ac.za

Chapter 7

Dissertation analysis

7.1 Statement on hedging

In this section, we briefly extend our discussion to one of the most important part of model implementation, namely hedging. The main part of this study mainly focused on pricing, pricing methods and calibration; however a model whose purpose is for pricing cannot be separated from hedging (Kienitz and Wetterau, 2012, chapter 10). Hedging is equally important as pricing. While pricing focuses on whether a pricing model can capture the distribution of underlying asset price at maturity date of an option, hedging performance measures the tendency of a pricing model to adequately capture the dynamics of the evolution of the underlying price process. Hedging performances is one of the best indicators of the effectiveness of a model. In essence, one should be able to hedge against prices from the implemented model.

Recall that in any pricing model, the price is expressed in terms of the model's underlying variables. One then ought to compute the mathematical derivative of the option price with respect to the underlying instrument. A common approach widely used for hedging analysis is delta hedging (also referred to as dynamic hedging). This involves finding the risk sensitivities representing the option price to changes in underlying instruments, which can be used to forecast market risks. The essence of delta hedging involves constructing a financial portfolio that can be rebalanced to achieve a desired financial position, and for this reason it plays a crucial role in risk management.

Chapter 2 introduced interest rate derivatives as crucial tools used to transfer risks due to the fluctuation of interest rates in the risk management. The main purpose of these derivatives is to transfer and minimise risks due to stochastic evolution of interest rates by means of hedging. As we have seen from Chapter 1, the frequently traded interest rate products in the fixed income market includes forward rate agreement (FRA) and swaps contracts which make up the zero-coupon bonds market. Options on these contracts are caps/floors and swaptions

Stellenbosch University https://scholar.sun.ac.za

respectively. Due to the high market size of these options, they requires accurate and efficient pricing and hedging performance. At first glance, from a financial mathematics point of view, the prices of these derivatives should be affected by the same factors that affect FRA and swap rates. However, recent empirical studies have shown that there are some risk factors that affect caps/floors and swaptions prices but do not affect the underlying FRA and swap rates respectively (see Haitao and Feng, 2006). This means that bonds do not dominate the fixed income market.

As mentioned before, empirical analysis shows that models driven by a Brownian motion fail to reproduce sufficient return distributions. To reduce model risk, classes of models that incorporates some leptokurtic features are studied and implemented. Eberlein and Raible (1999) showed that, by using general Lévy processes, it is not only the fit of the distribution is improved;in addition, general Lévy processes give a more realistic picture of option price movements.

The paper of Björk *et al.* (1997), which pioneered the inclusion of jumps using a general semimartingale processes in HJM models, shows that in models that exhibits jumps it is no longer a case that interest rates derivatives can be perfectly hedged. This type of model feature is known as unspanned stochastic volatility (USV). This means volatility risk cannot be lessened by using zero-coupon bonds. Hence, this makes bond market to be approximately complete. This means that interest rate models driven by Lévy processes create USV, which makes it difficult to hedge caps/floors and swaptions.

On the other hand, Eberlein *et al.* (2005) study the arbitrage and completeness problem in Lévy HJM models. They showed that a one-dimensional Lévy term structure model is complete provided that it has non-random coefficient. These findings agree with the original study by Björk *et al.* (1997), who showed that in a model with a continuum of financial securities there is a difference between completeness and the existence of equivalent martingale measure. Furthermore, they showed that in this arena, the existence of uniqueness equivalent martingale measure $\mathbb{Q}$ does not imply hedging of every contingent claim. This means a one-dimension Lévy HJM framework is approximately complete, i.e. hedging for any interest rate product is possible in an approximation sense (see Björk *et al.*, 1997, Section 6, p. 170). Hence one-dimensional Lévy HJM models are able to account for USV. Furthermore, in the bond market, perfect hedging is achieved by trading with an infinite number of bonds; this is however not possible, since the market provides finite number of bonds with different maturities.

We briefly discuss the lack of hedging literature for European interest rate contingent claim with a portfolio of zero-coupon bonds in Lévy HJM.

In Chapter 1, we outlined the fundamental distinction between the stock market and the bond market. In the bond market, hedging is a problematic because of infinite-dimensional vectors

Stellenbosch University https://scholar.sun.ac.za

of zero-coupon bonds, unlike in the stock market, where a hedging portfolio is constructed with a finite number of stocks. To make matters worse, hedging of Lévy term structure is a complicated subject as there are many factors contributing to risks. These include, the entire term structure, the jumps and the presence of an infinite number hedging instruments.

There is a very little literature about hedging in a classical HJM framework. This includes the notable work of Jarrow (1994), Rutkowski (1996) and Carmona and Tehranchi (2004). These all show that interest rates derivatives can be perfectly hedged by using portfolio of zero-coupon bonds. Numerically computation of hedging performance under HJM models are not yet implemented.

Even less is known about hedging in one-dimensional Lévy HJM "market" although "it" is known to be approximately complete, i.e. hedging is achieved by having a sequences of self-financing portfolios whose L^2 limits replicates the contingent claim. To the best of our knowledge of Lévy HJM models, there is only one paper that examines hedging of interest rate payer swaptions, a PhD thesis of Vandaele (2010). Vandaele (2010) numerically compared the performance of delta-hedging and quadratic hedging in a one-dimensional Lévy HJM.

There is a vast amount of literature on USV in term structure models to such an extent that there are conflicting views as to whether zero-coupon bonds span the fixed income market. The presence of USV seems to have significant implications for an interest rate model. However, this is not clear as several authors have positive arguments regarding USV while others are indifferent. Because implied volatilities for caps and floors are more volatile than the implied volatilities for swaptions, this makes the presence of USV has a minor effect on hedging swaptions, while it has serious implication for hedging caps/floors.

7.2 Limitations

This study was limited to the discussion of the driving process L which is a one dimensional processes generated by a generalised hyperbolic (GH) distributions. GH distributions incorporate many distributions widely used for financial modelling such as normal inverse Gaussian (NIG) and hyperbolic distribution. GH distributions offer more flexibilities to the driving process and we did not consider higher-dimensional driving processes. ”Note that it would not be appropriate to classify a model driven by a one-dimensional Lévy process as a one-factor model, since the driving Lévy process itself is already a high-dimensional object“, (Eberlein and Kluge, 2006, p. 22).

Stellenbosch University https://scholar.sun.ac.za

Chapter 8

Conclusion

In this dissertation we have studied and implemented the Heath–Jarrow–Morton (HJM) models driven by Lévy processes, namely Brownian motion (BM) and the generalised hyperbolic (GH) motion. The use of the GH motion is supported by statistical analyse showing that GH distributions give a clear description of financial data. This study arose from a two-fold problem: (1) empirical studies show that models driven by the BM do not reflect realities in bond markets, and (2) the great expansion of fixed income markets. Recent studies have tried to generalise HJM models. The cornerstone of this study was to analyse models driven by a Brownian motion and by a generalised hyperbolic process, implement the models in MATLAB, investigate the numerical valuation methods, and calibrate the models to the quoted South African market data for ATM caps/floors.

The study began with a brief review of a classical HJM framework for no-arbitrage. We also discussed the form of forward volatility structure that let short-rate processes to follow Markov processes.

We have opted to work with non-homogeneous Lévy processes in the driving process because of the flexibility they offer to the model. These include the possibility of derivation of analytical formulae. The main reason for preferring non-homogeneous over homogeneous Lévy processes is that the change of measure is vital as this avoids the calculations of joint distribution which is computationally expensive. We are aware that joint density was derived by Eberlein and Raible (1999), but its numerical evaluation is highly demanding. Furthermore, under new measures non-homogeneous Lévy processes retain their properties. That is to say that under the new martingale measure the driving process is still a process with independent increments and absolutely continuous characteristics (PIIAC).

We have chosen a model driven by a generalised hyperbolic Lévy process because its class of distribution is quite flexible so that we do not need to explore higher-dimensional driving processes. The GH distribution is also rich in structure in that it contains many other distri-

Stellenbosch University https://scholar.sun.ac.za

butions, such as the normal inverse Gaussian and hyperbolic distributions. Although it is a well-known that a three-factor model describes the patterns of the yield much better than a one-factor model, in our context of the driving process one cannot classify Lévy HJM model as in one-factor model because the driving process is already a high-dimensional entity, (see Eberlein and Kluge (2006)). Hence a one-dimensional driving process suffices for our analysis. Other than that, we are also very much aware of the Lévy term structure models driven by the multivariate Lévy process, (see Raible, 2000, Chapter 7).

The central concern is to value interest rate derivatives. Since no closed-form solutions exist, we have to use numerical approximations. We have discussed and implemented several numerical valuation methods. The first is the Monte Carlo method. It can be applied even if the change of measure is not required. The main challenge with this method is that the driving process is a high-dimensional entity and application of the Monte Carlo method entails computational complexity and is time consuming. Efficient valuation originated from the algorithms presented by Raible (2000), who used the bilateral Laplace transform, and the extension of his ideas by Eberlein and Kluge (2006). Among the methods discussed, we have implemented two methods new to the literature of Lévy term structure models, namely the COS method by Fang and Oosterlee (2008) and the fractional fast Fourier transform (FrFT) by Chourdakis (2004). Our numerical results for comparison show that the COS method and FrFT are very competitive. The COS method converges fast and with an error bound to the power of 10^{-16}. The main advantage of the COS method is that we do not have to worry about the dampening parameter which is proven to causes sensitivities to the option value. The only errors involved in this method arose from the density approximation and the truncation of the integration range. Since we are interested in the model calibration of interest rate derivatives which can have maturity of say three-months, it is shown that the COS method is not capable of handling options with a short times to maturity. To circumvent this, we proposed the integration range $[a, b]$ to be a function of maturity and we found $[-2\sqrt{T}, 2\sqrt{T}]$ to provide more stable results.

To prove the validation of a pricing model, at first glance, it must be calibrated to liquid data. In general, calibration for the term structure model is quite a challenging procedure, due to the nature and behaviour of the yield curve. We have calibrated our models to South African interest rate caps. In both models driven by a Brownian motion and by a generalised hyperbolic motion we obtained a nearly exact fit; however we used two different approaches of optimisation (secant-Levenberg–Marquardt and simulated Annealing). This provides for the accuracy of calibration to a sufficient extent. Although simulated annealing took longer, it resulted in the most effective optimisation. To compare the fit, we considered some objective functions. Our results show that a model driven by GH outperform a model driven by a BM. This is in line with expected results.

Possible further investigation could be numerical implementation of delta-hedging for cap-

Stellenbosch University https://scholar.sun.ac.za

s/floors in the HJM framework driven by a generalised hyperbolic motion.

Stellenbosch University https://scholar.sun.ac.za

Appendices

Stellenbosch University https://scholar.sun.ac.za

Appendix A

Subsequent theories

A.1 Discounted bond prices in Equation 4.3.13

Recall from Equation 4.2.2 that the short rate process is given by

$$r(t) = \mu(t) + \int_0^t \sigma(s,t)\, dL_s^{\mathbb{Q}},$$

where $\mu(t) = f(0,t) + \frac{\hat{\sigma}^2}{2a^2}\left(1 - e^{-at}\right)^2$ and L is a non-homogeneous Lévy process.

The discounted bond prices are given by

$$\begin{aligned}
Z(t,T) &= \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_0^T r(s)\, ds\right)\middle|\mathcal{F}_t\right] \\
&= \exp\left(-\int_0^T \mu(s)\, ds\right) \times \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_0^T \left(\int_u^T \sigma(u,s)\, dL_u\right)\, ds\right)\middle|\mathcal{F}_t\right] \\
&= \exp\left(-\int_0^T \mu(s)\, ds\right) \times \mathbb{E}_{\mathbb{Q}}\left[\exp\left(\int_0^T S(u,T)\, dL_u\right)\middle|\mathcal{F}_t\right] \\
&= \exp\left(-\int_0^T \mu(s)\, ds\right) \times \mathbb{E}_{\mathbb{Q}}\left[\exp\left(\int_0^t S(u,T)\, dL_u\right) \times \exp\left(\int_t^T S(u,T)\, dL_u\right)\middle|\mathcal{F}_t\right] \\
&= \exp\left(\int_0^t S(u,T)\, dL_u - \int_0^T \mu(s)\, ds\right) \times \mathbb{E}_{\mathbb{Q}}\left[\exp\left(\int_t^T S(u,T)\, dL_u\right)\right] \\
&= \exp\left(\int_0^t S(u,T)\, dL_u\right) \times \frac{P(0,T)}{\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\int_0^t S(u,T)\, dL_u\right)\right]}.
\end{aligned} \tag{A.1.1}$$

The fourth equality above holds because $\int_0^{\cdot} S(u,T)\, dL_u$ is a process with independent increments, (see Example 3.6.2).

Stellenbosch University https://scholar.sun.ac.za

A.2 The fast Fourier transform

The integral in Equation 5.3.11 can be seen as a inverse Fourier transform of the function $H(u) = M_{X_T}(R - \mathrm{i}u)\hat{\Phi}(u + \mathrm{i}R)$. Hence we can apply fast Fourier transformation (FFT). FFT allows rapid calculation of the inverse Fourier transforms with a vectors of strikes easily and efficiently.

Let $H = (h_0, \cdots, h_{N-1})$ be N vector in $\mathbb{R}$. We aim at evaluating

$$\int_0^\infty e^{-\mathrm{i}xu} h(u) du \approx \sum_{j=0}^{N-1} e^{-\mathrm{i}\frac{2\pi}{N}kj} h_j, \tag{A.2.1}$$

where we impose a constraint that we carefully choose $\vec{H} = (h_j)_0^{N-1}$, i.e. a vector consisting of N function evaluations of the function h, at the points $\vec{u} = (u_j)_{j=0}^{N-1}$.

Let us assume we would like to evaluate option prices using the discrete FT (DFT) approach. Denote the DFT of $\vec{h}$ by

$$D_n(\vec{h}) = \sum_{j=0}^{N-1} \alpha_n^j h_j \quad \text{where} \quad \alpha_n = \exp\left(-\mathrm{i}\frac{2\pi}{N}n\right).$$

The output of the DFT is another vector given by $\vec{f} = (f_j)_0^{N-1}$. Carr and Madan (1999) showed that $\vec{f}$ contain option prices that correspond to the log-strike and are contained in a vector $\vec{x} = (x_j)_0^{N-1}$.

The corresponding inverse discrete Fourier transform is given by

$$\mathcal{F}^{-1}[D_n(\vec{h})] = \frac{1}{N}\sum_{n=0}^{N-1} \alpha_j^{-n} D_n(\vec{H}).$$

The computation of the vector $D_n(\vec{h})$ requires about N^2 and $N^2 - N$ multiplications and additions respectively. We can increase the speed of computation if we consider the order of arithmetic operation to be $\frac{1}{2}N\log_2 N$.

To calculate

$$\int_{\mathbb{R}} e^{-\mathrm{i}ux} H(u) du = \mathcal{F}^{-1}\left[H(u)\right]$$

we need to find a point M and step size Δu that truncate the infinite interval so that we can work with

$$\int_{-M}^{M} e^{-\mathrm{i}ux} H(u)\, du.$$

Typically we choose $M = 2N - 1$.

We then apply Simpson's rule to find an approximation of the integral with width length of Δu. Denotes the above integral by

$$I(x_j) = \sum_{n=-(N-1)}^{n=N-1} \eta_n e^{-\mathrm{i}x_j n\Delta u} H(n\Delta u), \text{ where } \eta_n = \frac{\Delta u}{3}\left(3 + (-1)^{n+1} - \delta_n\right),$$

Stellenbosch University https://scholar.sun.ac.za

where δ is a Kronecker delta defined by

$$\delta_n = \begin{cases} 1, & \text{for n=0} \\ 0, & \text{otherwise.} \end{cases}$$

$H(-u) = \overline{H(u)}$ because it contains Fourier transform of real-valued functions, and the complex conjugate eliminates the complex part of the integral. We can therefore only evaluate for the positive terms because $\int_{-c}^{c} H(u) + \overline{H(u)}\ du = \int_0^c 2\Re\{H(u)\}$. This means we write the integral as

$$I(x_j) = 2\Re\left\{\sum_{n=0}^{n=N-1} \eta_n e^{-\mathrm{i}x_j n\Delta u} H_n\right\}.$$

Where we choose $H_0 = \frac{H(0)}{2}$ and for $n = 1 \cdots N-1$ we choose $H_n = H(n\Delta u)$. Ultimately, we have the representation

$$\int_{\mathbb{R}} e^{-\mathrm{i}ux} H(u) du = \Re\left\{\sum_{n=0}^{n=N-1} \eta_n e^{-\mathrm{i}x_j n\Delta u} H_n\right\}.$$

For evenly spaced values of $x = (x_0, \cdots, x_{N-1})$ with the grid width of Δx, we choose $x_0 = \frac{-N}{2}\Delta x$, $x_{N-1} = (\frac{N}{2} - 1)\Delta x$ and $x_j = x_0 + j\Delta x$. Then

$$I(x_j) \approx \sum_{n=0}^{N-1} \eta_n e^{-\mathrm{i}(x_0 + j\Delta x)n\Delta u} H_n = \sum_{n=0}^{N-1} \eta_n e^{-\mathrm{i}nx_0\Delta u} e^{-\mathrm{i}jn\Delta x\Delta u} H_n.$$

Let $\Delta x\Delta u = \frac{2\pi}{N}$ then the integral approximation becomes

$$I(x_j) \approx \sum_{n=0}^{N-1} \left(e^{-\mathrm{i}\frac{2\pi}{N}j}\right)^n F_n, \quad F_n = \eta_n e^{-\mathrm{i}x_0 n\Delta u} H_n,$$

which is in the form of a discrete Fourier transform, and hence can be computed fairly easily.

A.3 Characteristic function

We show how to get a T-forward characteristic function of a random variable

$$X_T = \int_0^T S(s, U) - S(s, T)\, dL_s.$$

The underlying theory here is the Girsanov transformation state in Theorem 3.5.4 and the change of measure techniques discussed in Section 4.8.

$$\chi_{X_T}(u) = \mathbb{E}_{\mathbb{Q}^T}\left[\exp\left(\mathrm{i}uX_T\right)\right] = \exp\left(\int_0^T \psi_s^{\mathbb{Q}^T}\left(iu[S(s, U) - S(s, T)]\right) ds\right).$$

Stellenbosch University https://scholar.sun.ac.za

The Lévy triplets for $\psi_s^{\mathbb{Q}^T}$ are given in (4.8.2).

Then

$$\begin{aligned}
\psi_s^{\mathbb{Q}^T}(y) &= a_s^{\mathbb{Q}^T} y + \frac{1}{2} b_s^{\mathbb{Q}^T} y^2 + \int_{\mathbb{R}} (e^{xy} - 1 - xy)\, \nu_s^{\mathbb{Q}^T}(dx) \\
&= \left(a_s + b_s S(s,T) + \int_{\mathbb{R}} \left(e^{S(s,T)y} - 1\right) y\, \nu_s(dy) \right) y + \frac{1}{2} b_s y^2 + \int_{\mathbb{R}} (e^{xy} - 1 - xy)\, e^{S(s,T)y}\, \nu_s(dy) \\
&= a_s y + b_s S(s,T) y + \frac{1}{2} b_s y^2 + \int_{\mathbb{R}} \left(e^{[S(s,T)+y]x} - 1 - [y + S(s,T)]x\right) \nu_s(dx) \\
&\qquad - \int_{\mathbb{R}} \left(e^{S(s,T)x} - 1 - S(s,T)x\right) \nu_s(dx) \\
&= a_s[y + S(s,T)] + \frac{1}{2} b_s [y + S(s,T)]^2 + \int_{\mathbb{R}} \left(e^{[S(s,T)+y]x} - 1 - [y + S(s,T)]x\right) \nu_s(dx) \\
&\quad - a_s S(s,T) - \frac{1}{2} b_s ||S(s,T)||^2 - \int_{\mathbb{R}} \left(e^{S(s,T)x} - 1 - S(s,T)x\right) \nu_s(dx) \\
&= \psi_s(y + S(s,T)) - \psi_s(S(s,T)).
\end{aligned}$$

Hence

$$\psi_s^{\mathbb{Q}^T}(\mathrm{i}u(S(s,U) - S(s,T))) = \psi_s(\mathrm{i}uS(s,U) + (1 - \mathrm{i}u)S(s,T)) - \psi_s(S(s,T),$$

and thus

$$\chi_{X_T}(u) = \exp\left(\int_0^T \psi_s(\mathrm{i}uS(s,U) + (1 - \mathrm{i}u)S(s,T)) - \psi_s(S(s,T)\, ds \right).$$

A.4 Jamshidian decomposition

This section reviews the Jamshidian decomposition for options on coupon bonds. This decomposition was pioneered by Jamshidian (1989) for the Vasiček model and extended to any model by Brigo and Mercurio (2006). Moreover, Brigo and Mercurio (2006) showed that the Jamshidian decomposition is valid for the short-rate model as long as the price process of a zero-coupon bond is a decreasing function of the interest rate. Therefore, this decomposition is necessary for derivation of swaption pricing formulae.

Let $P(R,t,s)$ be the time t-price of a zero-coupon bond prevailing or maturing at time $s > t$ and R be a random variable. Using the same notation as introduced in Section 2.3.2, Jamshidian (1989) showed that the pay-off

$$\max\left\{ \sum_{i=1}^{n} c_i P(R,T,s_i) - \kappa, 0 \right\}$$

Stellenbosch University https://scholar.sun.ac.za

can be decomposed as

$$\sum_{i=1}^{n} c_i \max\{P(R,T,s_i) - \kappa_i, 0\},$$

where $\kappa_i = P(K,T,s_i)$ and K is the unique solution to the equation

$$\sum c_i P(K,T,s_i) - K = 0. \tag{A.4.1}$$

A.5 Proof: Theorem 6.1.1

Let $\Phi^i(x) = (K_i - f(T,T_i,x))^+$, for $i = 1, 2, \cdots, n$. Φ^i are generally not integrable, hence to enforce the degree of integrability, we dampened them and call $g^i(x) = e^{Rx}\Phi^i(x)$. Hence one can show that $g^i(x)$ satisfies all conditions postulated in Condition 5.3.3.

Recall from (5.3.10) that $\hat{\Phi}^i(\mathrm{i}R - u) = \hat{g}^i(-u)$.

$$\hat{\Phi}^i(\mathrm{i}R-u) = \hat{g}^i(-u) = \int_{\mathbb{R}} \Phi^i(x) e^{(-\mathrm{i}u-R)x}\, dx = \int_{-\infty}^{K} \left(K_i - D_T^{T_i} e^{R(T,T_i)x}\right) e^{(-\mathrm{i}u-R)x}\, dx$$

$$= K_i e^{(-\mathrm{i}u-R)K} \int_{-\infty}^{0} e^{(-\mathrm{i}u-R)K}\, dx - D_T^{T_i} e^{(R(T,T_i)-\mathrm{i}u-R)K} \int_{-\infty}^{0} e^{(R(T,T_i)-\mathrm{i}u-R)x}\, dx.$$

Now change variables, i.e. $t_1 = e^{x_1}$ and $t_2 = e^{x_2}$, where $x_1 = -\mathrm{i}u - R$ and $x_2 = R(T,T_i) - \mathrm{i}u - R$. Also recall the basic results about beta function

$$\int_0^1 t^{x-1}(1-t)^{y-1} = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Then,

$$\hat{\Phi}^i(\mathrm{i}R-u) = K_i e^{(-\mathrm{i}u-R)K} \int_0^t t^{-\mathrm{i}u-R-1}\, dt - D_T^{T_i} e^{(R(T,T_i)-\mathrm{i}u-R)K} \int_0^t t^{R(T,T_i)-\mathrm{i}u-R-1}\, dt$$

$$= K_i e^{(-\mathrm{i}u-R)K} \left(\int_0^1 t^{x_1-1}\, dt - \int_0^1 t^{x_2-1}\, dt \right) = K_i e^{(-\mathrm{i}u-R)K} \left(\frac{1}{x_1} - \frac{1}{x_2} \right).$$

The characteristic function of X_{T_n} has the same form as in Equation 5.3.2, defined by

$$\chi_{X_{T_n}}(u) = \exp\left(\int_0^T \psi_s\left(\mathrm{i}uS(s,T_n) + (1-\mathrm{i}u)S(s,T)\right) - \psi_s\left(S(s,T)\right) ds \right).$$

$\square$

Stellenbosch University https://scholar.sun.ac.za

Appendix B

MATLAB codes

This section includes the MATLAB codes. As stated before, we have used n-point Gauss–Legendre for integration. The scrips for Gauss–Legendre are taken from (Kiusalaas, 2010, chapter 6). One can also use MATLAB's built-in *integral* function.

B.1 Fourier integration method

```
function price = Fourier_method(Driver,PU,PT,U,T,K,params,R,L)
% Fourier_method: calculate the price for  option on ZCB
% Inputs: PU=> U-bond
%        : PT=> T-Bond
%        : U=> Bond maturity
%        : T=> Option maturity
%        : K=> Option strike price
%        : L=> Integration trancation range
%        : Parameters=>[k,lambda,alpha,beta,Δ,mu]
% Outputs: Price=Call if R>1
%                 =Put  if R<0
%
% References: Eberlein, Glau and papapantoleon: Analysis of Fourier
%             Transform valuation formulas and applications

% By Mesias Alfeus: Stellenbosch University
%*******************************************************************
% model parameters
sig=params(1);a=params(2);lambda=params(3);alpha=params(4);
beta1=params(5); Δ=params(6);mu=params(7);
   %log moment generating function for the drivers
     function y=lmgf(u,params)
```

Stellenbosch University https://scholar.sun.ac.za


```
        sig=params(1);a=params(2);lambda=params(3);alpha=params(4);
beta1=params(5); Δ=params(6);mu=params(7);
         if strcmp(Driver, 'BM')
        y=0.5*u.*u.*T;
         else
        y=mu.*u.*T+0.5*lambda.*T*(log(alpha^2-beta1^2)-log(alpha^2-(beta1+u).^2))+...
          log(besselk(lambda,Δ*sqrt(alpha^2-(beta1+u).^2)).^T)-...
          log(besselk(lambda,Δ*sqrt(alpha^2-beta1^2)).^T);
         end
    end

 % Bond volatility structure
function S=Vasicekvol(t,T,sig,a)
        S=(sig/a)*(1-exp(-a.*(T-t)));

end

% Moment generating function a driving process X under
% forward martingale measure
 function y=mgf(u,T,U,sig,a,params)

y = exp( gaussQuad(@(t) ...
    (lmgf(u.*Vasicekvol(t,U,sig,a)+(1-u).*Vasicekvol(t,T,sig,a),params)-...
lmgf(Vasicekvol(t,T,sig,a),params)),0,T,7) );
 end
% Fourier transform of dampened payoff function
    function y=Laplace(z)
        y = K.^(1+complex(0,1)*z)/(complex(0,1)*z*(1+complex(0,1)*z));
    end

% Bond  Deterministic  part
D=(PU/PT).*exp(gaussQuad ( @(t) (lmgf(Vasicekvol(t,T,sig,a),params)-...
lmgf(Vasicekvol(t,U,sig,a),params)),0,T,7));

% Option value
price=PT.*real(gaussQuad(@(u) ...
    D.^(complex(R,-u)).*mgf(complex(R,-u),T,U,sig,a,params)...
.*Laplace(complex(u,R))./(2*pi), -L,L,8));


end
```

B.2 Fast Fourier transform method

Stellenbosch University https://scholar.sun.ac.za


```
function option=FFT_method(Driver,PU,PT,U,T,K,params,N,R)
% FFT_method: calculate the price for option on ZCB
% Inputs: PU=> U-bond
%        : PT=> T-Bond
%        : U=> Bond maturity
%        : T=> Option maturity
%        : K=> Option strike price
%        : Parameters=>[a lambda alpha beta Δ]
%        : N points
% Outputs: Price=Call if R<0
%                =Put  if R>1
%
% References: Raible (2000): Levy processes in Finance, theory, Numerics
%                             and impirical Facts (Section 3.6)

% By Mesias Alfeus: Stellenbosch University
%***************************************************************************
% model parameters
sig=params(1);a=params(2);lambda=params(3);alpha=params(4);
beta1=params(5); Δ=params(6);mu=params(7);
%log moment generating function for the drivers
     function y=lmgf(u)

         if strcmp(Driver, 'BM')
        y=0.5*u.*u.*T;
         else
        y=mu.*u.*T+0.5*lambda.*T*(log(alpha^2-beta1^2)-log(alpha^2-(beta1+u).^2))+...
          log(besselk(lambda,Δ*sqrt(alpha^2-(beta1+u).^2)).^T)-...
          log(besselk(lambda,Δ*sqrt(alpha^2-beta1^2)).^T);
         end
     end

% Bond Vol
 function S=Vasicekvol(t,T,sig,a)
        S=(sig/a)*(1-exp(-a.*(T-t)));
 end
% Characteristic function  of a driving process X under
% forward martingale measure
 function y=CF(u,T,U)
y = exp( gaussQuad(@(t)  (lmgf(1i*u.*Vasicekvol(t,U,sig,a)+(1-1i.*u).*...
    Vasicekvol(t,T,sig,a))-lmgf(Vasicekvol(t,T,sig,a))),0,T,7) );
 end
% FFT methods
n = 2^N;               % number of points in integration
```

Stellenbosch University https://scholar.sun.ac.za


```
du=0.25;               % Step size in integration
dx=(2*pi)/(du*n);      %spacing for  strike
gamma=(n*dx)/2;

u1=1:n;                % indices
x=-gamma+dx.*(u1-1);   % log strike level -gamma to gamma
j=1:n;

u=(j-1).*du;

%calculate the Laplace transform of the contract function K=1;
 function y=Laplace(v)
     %K=1
        y = 1./(v.*(1+v));
 end
% Convolution
Hn = CF(complex(-u,R),T,U).*Laplace(complex(R,u));
% calculate the sequence y
y=exp(-1i.*u.*gamma).*Hn;
kk=1:n;
% apply the  Simpson integration rule to y
y=(y./3).*(3 + (-1).^kk - ((kk-1)==0));
% perform the FFT to get Gk
Gk = fft(y);
% put option price Pk
Pk=real(du.*(PT.*exp(-R.*x)/pi).*Gk);

% Bond  Deterministic  part
D=(PU/PT).*exp(gaussQuad ( @(t)  (lmgf(Vasicekvol(t,T,sig,a))-...
 lmgf(Vasicekvol(t,U,sig,a))),0,T,7));
% get  strike prices
strikes= exp(log(D)-x);

% Then the price
prices=strikes.*Pk;
% Option value
option = interp1(strikes,prices,K);
end
```

B.3 Fractional fast Fourier transform method

```
function option=FrFT_method(Driver,PU,PT,U,T,K,params,PA,PB,N,R)
```

Stellenbosch University https://scholar.sun.ac.za


```
% FrFT_method: calculate the price for option on ZCB
% Inputs: PU=> U-bond
%        : PT=> T-Bond
%        : U=> Bond maturity
%        : T=> Option maturity
%        : K=> Option strike price
%        : Parameters=>[a lambda alpha beta Δ]
%        : N points
%        : PA upper bound for integration
%        : PB bound for log strike (-PB,PB)
% Outputs: Price=Call if R<0
%               =Put  if R>1
%
% References: Raible (2000): Levy processes in Finance, theory, Numerics
%                            and impirical Facts (Section 3.6)
%                          : Chourdakis (2004)
% By Mesias Alfeus: Stellenbosch University
%***************************************************************************
% model parameters
sig=params(1);a=params(2);lambda=params(3);alpha=params(4);
beta1=params(5); Δ=params(6);mu=params(7);
%log moment generating function for the drivers
     function y=lmgf(u)

   if strcmp(Driver, 'BM')
   y=0.5*u.*u.*T;
   else
  y=mu.*u.*T+0.5*lambda.*T.*(log(alpha^2-beta1^2)-log(alpha^2-(beta1+u).^2))+...
   log(besselk(lambda,Δ*sqrt(alpha^2-(beta1+u).^2)).^T)-...
   log(besselk(lambda,Δ*sqrt(alpha^2-beta1^2)).^T);
  end
  end

% Bond Vol
 function S=Vasicekvol(t,T,sig,a)
        S=(sig/a)*(1-exp(-a.*(T-t)));
 end
%calculate the Laplace transform of the contract function K=1;
 function y=Laplace(v)
     %K=1
        y = 1./(v.*(1+v));
 end
% Characteristic function  of a driving process X under
% forward martingale measure
 function y=CF(u,T,U)
```

Stellenbosch University https://scholar.sun.ac.za


```
  y = exp( gaussQuad(@(t) (lmgf(1i*u.*Vasicekvol(t,U,sig,a)+(1-1i.*u).*...
Vasicekvol(t,T,sig,a))-lmgf(Vasicekvol(t,T,sig,a))),0,T,7) );
 end
% FrFT methods

function f=frft(x,a1)
N1=size(x,2);
e1=exp(-pi*1i*a1*(0:(N1-1)).^2);
e2=exp(pi*1i*a1*(N1:-1:1).^2);
z1=[x.*e1,zeros(1,N1)];
z2=[1./e1,e2];
fz1=fft(z1);
fz2=fft(z2);
fz=fz1.*fz2;
ifz=ifft(fz);
f=e1.*ifz(1,1:N1);
end

n = 2^N;                % number of points in integration
du=PA/n;                % Step size in integration
dx=(2*PB)/n;            % spacing for  strike
a1=(du.*dx)/(2*pi);     % fractional parameter
gamma=-PB;              % first grid


x=gamma+dx.*(0:n-1);    % log strike level -gamma to gamma
u=(0:n-1).*du; %(u)

% Convolution
Hn = CF(complex(-u,R),T,U).*Laplace(complex(R,u));
% Integrand
y=exp(-1i.*u.*gamma).*Hn.*du;
kk=1:n;
% apply the  Simpson integration rule to y
y=(y./3).*(3 + (-1).^kk - ((kk-1)==0));
% perform the frft to get Gk
Gk =frft(y,a1); % a=alpha
% put option price Pk
Pk=real((PT.*exp(-R.*x)/pi).*Gk);

% Bond  Deterministic  part
D=(PU/PT).*exp(gaussQuad ( @(t) (lmgf(Vasicekvol(t,T,sig,a))-...
 lmgf(Vasicekvol(t,U,sig,a))),0,T,7));
% get  strike prices
strikes= exp(log(D)-x);
```

Stellenbosch University https://scholar.sun.ac.za


```

% Then the price
prices=strikes.*Pk;
% Option value
option = interp1(strikes,prices,K);
end
```

B.4 Cosine method

```
function [call,put]=Cosine_method(Driver,PU,PT,U,T,K,params,N)
% Cosine_method: Evaluate call and put value using COS method
% inputs: Driver=> 'BM' or 'GH'
%         : K    =>    strike
%         : PU-Bond maturing at time U
%         : PT-Bond maturing at time T
%         : U-Bond maturity
%         : T -Option maturity
%         : N for the number of points
%         : params=> [k lambda alpha beta Δ];

% outputs: call option
%          : put option

% Main reference: F. Fang and C.W. Oosterlee (2008) " A novel pricing for
%                  European options based on Furier-Cousine series expansions"
%
% 2014 Mesias Alfeus
%*********************************************************************************
% model parameters
sig=params(1);k=params(2);lambda=params(3);alpha=params(4);
beta1=params(5); Δ=params(6);mu=params(7);
%log moment generating function for the drivers
 function y=lmgf(u)
 if strcmp(Driver, 'BM')
  y=0.5*u.*u.*T;
 else
 y=mu.*u.*T+0.5*lambda.*T*(log(alpha^2-beta1^2)-log(alpha^2-(beta1+u).^2))+...
  log(besselk(lambda,Δ*sqrt(alpha^2-(beta1+u).^2)).^T)-...
  log(besselk(lambda,Δ*sqrt(alpha^2-beta1^2)).^T);
  end
  end

```

Stellenbosch University https://scholar.sun.ac.za


```
% Bond Vol
 function S=Vasicekvol(t,T,sig,k)
        S=(sig/k)*(1-exp(-k.*(T-t)));
 end
% Coeffient for put option
function y=vk_put(k,b,a,Strike)
[Y1, Y2]=coef(k,a,0,a,b);
y=(2./(b-a)).*(Y2-Y1)*diag(Strike);
end

function [Y1, Y2]=coef(k,c,d,a,b)
Y1=(exp(d)-exp(c));
Y2=(d-c);
V=double((b-a)./(k.*pi).*(sin(k.*pi.*(d-a)./(b-a))-sin(k.*pi.*(c-a)./(b-a))));
Y2(2:end,:)=V(2:end,:);
Y11=1./(1+(k.*pi./(b-a)).^2);
Y12=exp(d).*cos(k.*pi.*(d-a)./(b-a))-exp(c).*cos(k.*pi.*(c-a)./(b-a));
Y13=k.*pi./(b-a).*(exp(d).*sin(k.*pi.*(d-a)./(b-a))-exp(c).*...
sin(k.*pi.*(c-a)./(b-a)));
V=Y11.*(Y12+Y13);
Y1(2:end,:)=V(2:end,:);
end
% Characteristic function  of a driving process X under
% forward martingale measure
 function y=CF(u,T,U,sig,k)
  y = exp( gaussQuad(@(t) (lmgf(1i*u.*Vasicekvol(t,U,sig,k)+(1-1i*u).*...
Vasicekvol(t,T,sig,k))-lmgf(Vasicekvol(t,T,sig,k))),0,T,7) );
 end

n=2^N;
L=2; % our suggestions
NK=size(K,1); % number of strikes


% Bond  Deterministic  part
D=(PU/PT).*exp(gaussQuad ( @(t) (lmgf(Vasicekvol(t,T,sig,k),params)-...
lmgf(Vasicekvol(t,U,sig,k),params)),0,T,7));

% The center
x=repmat(double(log(D./K))',n,1);
a=x-L*sqrt(T);b=x+L*sqrt(T); % Our choice

kn=repmat((0:n-1)',1,NK);
% Coeffiecient for put
vkp=@(x) vk_put(x,b,a,K);
```

Stellenbosch University https://scholar.sun.ac.za


```

% calculating the A_k (3.41)
Ak=double(real(CF(kn.*pi./(b-a),T,U,sig,k).*exp(1i.*pi.*kn.*x./(b-a)).*...
exp(-1i.*(pi.*kn.*a./(b-a)))));

Vk=vkp(kn); % Coeffient
put=double(PT.*(sum(Ak.*Vk)-0.5*(Ak(1,:).*Vk(1,:))))'; % put option

% Put-call parity
call=put+PU-K*PT;
end
```

B.5 Swaption pricing

```
function option=Fourier_Swaption(Driver,P,T,kappa,tenor,params,R,L)
% Fourier_Swaption: calculate the price for swaption
% Inputs: P=> vector of initial zero-coupon bonds
%        : T=> vector of maturity times
%        : kappa=> ATM strike
%        : tenor=> tenor
%        : Params=>[sig, a lambda alpha beta Δ]
%        : L=> integration range
% Outputs: Price=Reciever if R>1
%                =Payer    if R<0
%
% References: Eberlein and Kluge (2004):Exact pricing formulae for caps and
%                                       swaptions in a Levy term structure ...
    model

% By Mesias Alfeus: Stellenbosch University
%*************************************************************************

% model parameters
sig=params(1);a=params(2);lambda=params(3);alpha=params(4);
beta1=params(5); Δ=params(6);
%log moment generating function for the drivers
     function y=lmgf(u)
         if strcmp(Driver, 'BM')
        y=0.5*u.*u.*T;
         else
        y=0.5*lambda*T*(log(alpha^2-beta1^2)-log(alpha^2-(beta1+u).^2))+...
          log(besselk(lambda,Δ*sqrt(alpha^2-(beta1+u).^2)).^T)-...
```

Stellenbosch University https://scholar.sun.ac.za


```
             log(besselk(lambda,Δ*sqrt(alpha^2-beta1^2)).^T);
            end
        end

% Bond Volatility structure
 function S=Vasicekvol(t,T,sig,a)
        S=(sig/a)*(1-exp(-a.*(T-t)));
 end

% moment generating function for X_T
    function y=mgf(u,T,U)

y = exp(gaussQuad(@(t) ...
    (lmgf(u.*Vasicekvol(t,U,sig,k)+(1-u).*Vasicekvol(t,T,sig,a))...
     -lmgf(Vasicekvol(t,T,sig,a))),0,T,7) );
    end
% Coupon
N=length(P)-1;
C=ones(N,1)*Kappa*tenor;
C(N,1)=1+Kappa*tenor;
% Deterministic part
D=zeros(N,1);
for i=1:N
D(i)=(P(i)/P(1)).*C(i).*exp(gaussQuad( @(t) (lmgf(Vasicekvol(t,T(1),sig,a))-...
     lmgf(Vasicekvol(t,T(i),sig,a))),0,T(1),7));
end

B=zeros(N,1);
for i=1:N
    B(i)=gaussQuad( @(t) exp(-k*t),0,T(i),7)./gaussQuad( @(t) ...
       exp(-k*t),0,T(N),7);
end

function y=strkefunction(D,B,x)
y=sum(D.*exp(B.*x))-1;
end


% Get a Strike price
K=fsolve(@(x) strkefunction(D,B,x),0.01);


% The Bilateral Laplace
function f=laplace(D,B,K,u)
f=exp(u.*K).*(sum(D.*exp(B.*K).*((-1)./(B+u)))+1./u);
```

Stellenbosch University https://scholar.sun.ac.za


```
end

% Final value
option=P(1).*gaussQuad( @(u) real(laplace(D,B,K,complex(R,u))...
      .*mgf(complex(-R,-u),T(1),T(N)).*(1/(2*pi)) ), -L,L,7);
end
```

Stellenbosch University https://scholar.sun.ac.za

List of references

Bingham, N. and Kiesel, R. (2004). *Risk-neutral valuation: Pricing and hedging of financial derivatives*. Springer Finance, London. 10

BIS (2013). Statistical release: OTC derivatives statistics at end - June 2013. Available online: http://www.bis.org/publ/otc_hy1311.pdf. 1

Björk, T. (1995). On the term structure of discontinuous interest rates. *Surveys in Industrial and Applied Mathematics*, vol. 2, no. 4, pp. 626–657. 4

Björk, T. (2009). *Arbitrage theory in continuous time*. Oxford University Press, Oxford. 10, 39

Björk, T., Di Masi, G., Kabanov, Y. and Runggaldier, W. (1997). Toward a general theory of bond markets. *Finance and Stochastics*, vol. 1, no. 1, pp. 141–174. 4, 6, 8, 38, 99

Brigo, D. and Mercurio, F. (2006). *Interest rate models–theory and practice: With smile, inflation and credit*. Springer Finance, Berlin. 10, 14, 108

Carmona, R. and Tehranchi, M. (2004). A characterization of hedging portfolios for interest rate contingent claims. *The Annals of Applied Probability*, vol. 14, no. 3, pp. 1267–1294. 100

Carr, P. and Madan, B. (1999). Option valuation using the fast Fourier transform. *Journal of Computational Finance*, vol. 2, no. 4, pp. 61–73. 7, 60, 62, 63, 106

Carverhill, A. (1994). When is short rate markovian? *Mathematical Finance*, vol. 4, no. 4, pp. 305–312. 43, 50

Chourdakis, K. (2004). Option pricing using the fractional FFT. *Journal of Computational Finance*, vol. 8, no. 2, pp. 1–18. 7, 71, 102

Cont, R. and Tankov, P. (2004). *Financial modeling with jump processes*. Chapman and Hall/CRC Press, Oxford. 19, 25, 29, 33, 36, 37, 44

Cox, J., Ingersoll, J. and Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica*, vol. 53, no. 2, pp. 385–407. 2

Stellenbosch University https://scholar.sun.ac.za

Das, S. (1998). *Poisson-Gaussian Processes and the Bond Market*. NBER working paper series. National Bureau of Economic Research.
Available at: http://books.google.com.na/books?id=NWAkYAAACAAJ 4

Deng, D. and Chong, S. (2011). Efficient option pricing methods based on Fourier series expansions. *Journal of Mathematical Research & Exposition*, vol. 31, no. 1, pp. 12–22. 72

Eberlein, E. (2001). Application of generalized hyperbolic Lévy motion to finance. In: O.E. Barndorff-Nielsen, T.M. and Resnick, S. (eds.), *Lévy Processes: Theory and Application*, pp. 319–337. Birkhauser, Boston. 6, 19, 26, 27, 28, 92

Eberlein, E., Glau, K. and Papapantoleon, A. (2010). Analysis of Fourier transform valuation formulas and applications. *Applied Mathematical Finance*, vol. 17, no. 2, pp. 211–240. 7, 60, 64, 65

Eberlein, E., Jacob, J. and Raible, S. (2005). Lévy term structure models: No-arbitrage and completeness. *Finance and Stochastics*, vol. 9, pp. 67–88. 5, 6, 8, 38, 47, 99

Eberlein, E. and Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli*, vol. 1, no. 3, pp. 281–299. 27

Eberlein, E. and Kluge, W. (2006). Exact pricing formulae for caps and swaptions in a Lévy term structure model. *Journal of Computational Finance*, vol. 9, no. 2, pp. 99–125. 7, 49, 65, 85, 100, 102

Eberlein, E. and Prause, K. (). The generalized hyperbolic model: Financial derivatives and risk measures. Mathematical Finance- Bachelier Congress 2000, Springer. 2, 26, 28

Eberlein, E. and Raible, S. (1999). Term structure models driven by general Lévy processes. *Mathematical Finance*, vol. 9, no. 1, pp. 31–53. 2, 4, 5, 6, 19, 38, 43, 44, 46, 50, 52, 53, 76, 99, 101

Fang, F. and Oosterlee, C. (2008). A novel pricing method for european options based on Fourier-cosine expansions. *Numerische Mathematik*, vol. 114, pp. 27–62. 7, 72, 74, 102

Filipović, D. and Tappe, S. (2008). Existence of Lévy term structure models. *Finance and Stochastics*, vol. 12, pp. 83–115. 38

Geman, H., El Karoui, N. and Rochet, J. (1995). Change of numeraire, changes of probability measure and pricing of options. *Journal of Applied Probability*, vol. 4, no. 32, pp. 443–458. 54

Glasserman, P. (2003). *Monte Carlo methods in financial engineering*. Springer-Verlag, New York. 7

Haitao, L. and Feng, Z. (2006). Unspanned stochastic volatility: Evedence from hedging interest rate derivatives. *The Journal of Finance*, vol. 61, no. 1, pp. 341–378. 99

Heath, D., Jarrow, R. and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, vol. 60, no. 1, pp. 77–105. 3, 38, 40

Stellenbosch University https://scholar.sun.ac.za

James, J. and Webber, N. (2000). *Interest rate modelling*. John Wiley, New York. 10

Jamshidian, F. (1989). An exact bond option formula. *Journal of Finance*, vol. 44, no. 1, pp. 205–209. 84, 108

Jarrow, R. (1994). Delta, gamma and bucket hedging of interest rate derivatives. *Applied Mathematical Finance*, vol. 1, no. 1, pp. 21–48. 100

Jarrow, R. and Madan, D. (1995). Option pricing using the term structure of interest rates to hedge systematic discontinuities in asset returns. *Mathematical Finance*, vol. 5, no. 4, pp. 311–336. 6

Kienitz, J. and Wetterau, D. (2012). *Financial modelling: Theory, implementation and practice with MATLAB source*. Wiley Finance, Oxford. 90, 91, 98

Kirkpatrick, S., Gelatt, C. and Vecchi, M. (1983). Optiminization by simulated annealing. *Science*, vol. 1, no. 23, pp. 671–680. 90

Kiusalaas, J. (2010). *Numerical methods in Engineering with MATLAB*. Cambridge University Press, Cambridge. 110

Kluge, W. (2005). *Time-inhomogeneous Lévy processes in interest rate and credit risk models*. PhD, University of Freiburg. 4, 5, 6, 7, 33, 38, 45, 50, 62

Kuan, G. and Webber, N. (2001). Modelling interest rates with Lévy processes. Faculty of Economy, University of Piemonte Orientale, http://oldweb.eco.unipmn.it/eventi/irm/webber.pdf. 7, 65

Lewis, A. (2001). A simple option formula for general jump-diffusion and other exponential Lévy processes. OptionCity.net publications: http://optioncity.net/pubs/ExpLevy.pdf. 60

Musiela, M. and Rutkowski, M. (2004). *Martingale methods in financial modelling*. Springer, Sydney. 10, 14, 18, 58

Raible, S. (2000). *Lévy processes in finance: Theory, numerics, and empirical facts*. PhD, University of Freiburg. 2, 4, 5, 6, 7, 30, 38, 60, 62, 63, 65, 68, 70, 81, 84, 102

Ritchken, P. and Sankarasubramanian, L. (1995). Volatility structures of forward rates and the dynamics of the term structure. *Mathematical Finance*, vol. 5, no. 1, pp. 55–72. 43, 50

Rutkowski, M. (1996). Valuation and hedging of contingent claims in the HJM models with deterministic volatilities. *Applied Mathematical Finance*, vol. 3, no. 3, pp. 237–267. 100

Sato, K. (2001). Basic results on Lévy processes. In: O.E. Barndorff-Nielsen, T.M. and Resnick, S. (eds.), *Lévy Processes: Theory and Application*, pp. 3–37. Birkhauser, Boston. 19, 20, 21, 23

Schoutens, W. (2003). *Lévy processes in finance: Pricing financial derivatives*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., New York. 19

Stellenbosch University https://scholar.sun.ac.za

Shirakawa, H. (1991). Interest rate option pricing with Poisson-Gaussian forward curve processes. *Mathematical Finance*, vol. 1, no. 4, pp. 77–94. 4, 6

Vandaele, N. (2010). *Quadratic hedging in finance and insurance*. PhD, University of Gent. 8, 100

Vasiček, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, vol. 5, no. 2, pp. 177–188. 2